Scaling in Computer Network Traffic

Darryl Veitch

http://www.cubinlab.ee.mu.oz/∼darryl

Department of Electrical & Electronic Engineering
The University of Melbourne

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **07 JAN 2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Scaling in Computer Network Traffic**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical & Electronic Engineering The University of Melbourne**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001750, Wavelets and Multifractal Analysis (WAMA) Workshop held on 19-31 July 2004., The original document contains color images.

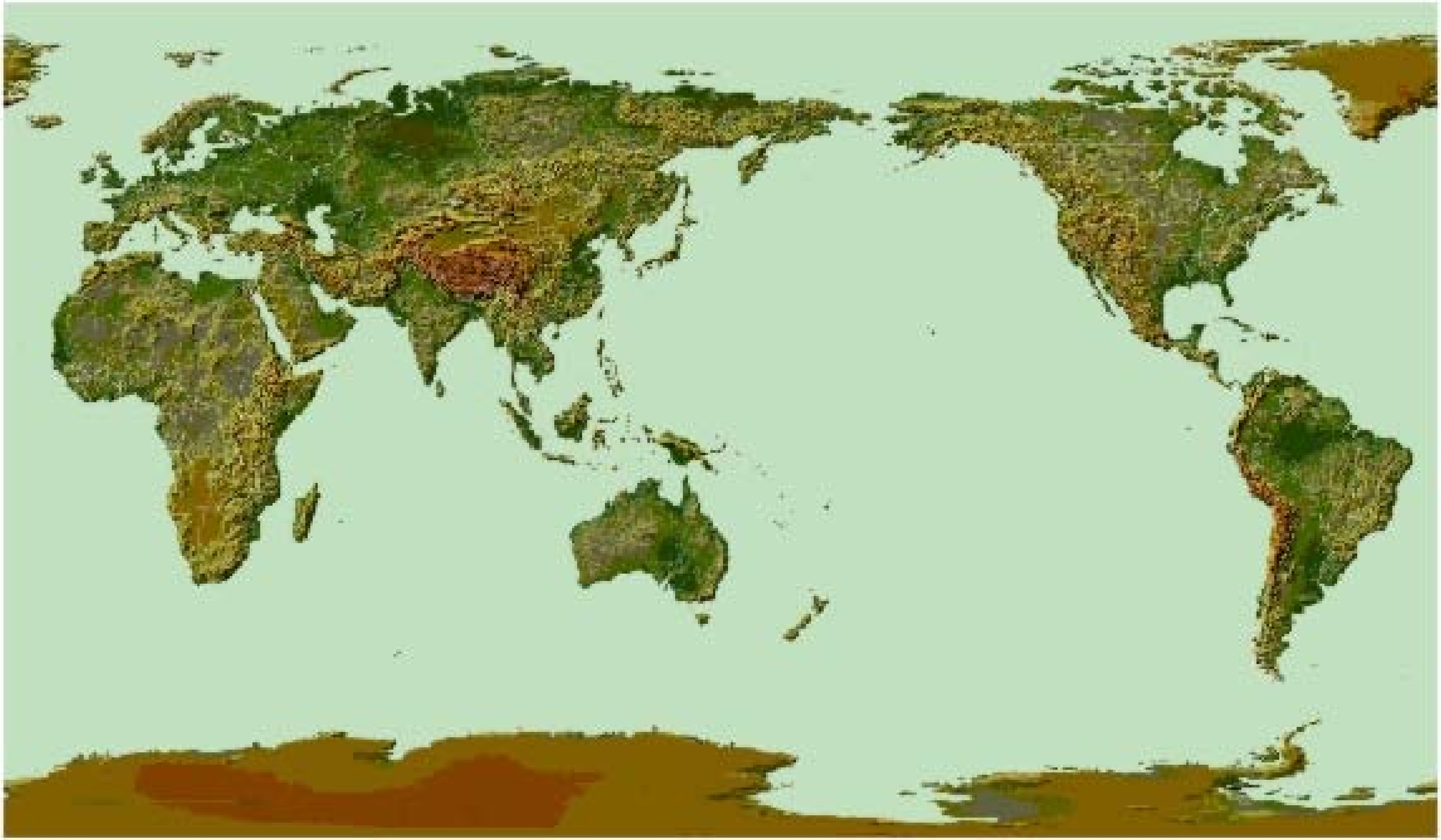
14. ABSTRACT

15. SUBJECT TERMS

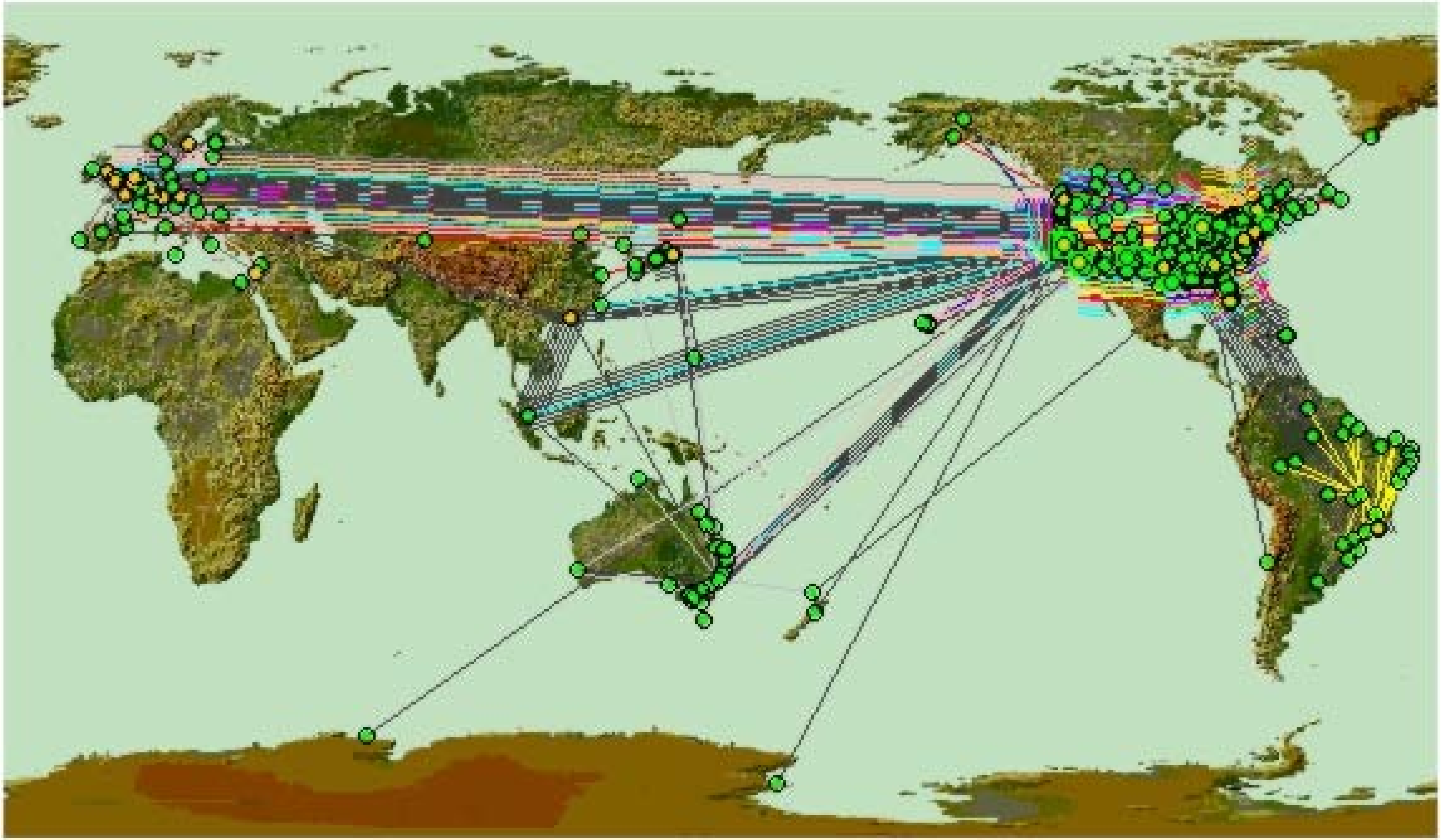
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**57**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Networks ...



Networks ... Connect



Telecommunications Networks, Traffic, & Engineering

Networks: A Deep Hierarchy of Systems

Tele–Traffic: A Turbulent River of Myriad Data Sources

Engineering: Traffic over a Network: Designing, Managing, Optimising

Networks: A Deep Hierarchy of Systems

- Connectivity (node organisation, placement, topology).
- Physical layer technology (devices, timing, coding, error recovery..).
- Circuit *vs* Packet Switched Paradigms.
- Connection-oriented *vs* connectionless philosophy.
- Protocol hierarchy and encapsulation, eg: Ethernet [IP [TCP[HTTP]]].
- Operation (routing, signaling, admission control, congestion control).
- Inter-Networking (Autonomous systems, domain routing, addressing, gateways, protocol translation).

Tele–Traffic: A Turbulent River of Myriad Data Sources

- ‘Geographic’ Complexity
  - At network edge (distribution of sources, destinations, nodes)
  - Internally (multiplexing and de-multiplexing of streams)
- Offered Traffic Complexity
  - User ‘sessions’ (durations, arrivals, number of clicks..)
  - Applications used (browser, napster..)
  - Underlying protocols (TCP, HTTP..)
  - Underlying data objects (files, video, audio..)
  - Shaping by network elements
- Temporal Complexity:
  - Human driven (diurnal, coffee breaks, think times..)
  - Source driven (real-time constraints, file sizes..)
  - Network driven (topology, protocols..)
  - Technology driven (link rates, switching rates)

$\longrightarrow$ Time scale rich: ns to months, 1kbps to Terabits/s

$\longrightarrow$ Burstiness: temporal (scaling), amplitude (non-G), spatial (capacity diversity)

TeleTraffic Engineering: the traditional program

- Collect traffic measurements and measure characteristics
- Select traffic model
- Solve switch performance problems
- Solve network performance problems
- Design network to given quality of service at minimal cost
- Solve admission control problems
- Run network
- Solve congestion control and routing problems
- Improve network performance

$\longleftarrow$ Iterate until paradigm shift

This ‘scientific method’ approach begins with measurements...

Papers between 1966 and 1987 (P. F. Pawlita, ITC-12, Italy)

- Queueing theory: several thousand
- Traffic measurement: around 50.

Tele–Traffic Today

Measurement Practice

- Recognition of measurement, not just routine quantification, but discovery.
- Widespread monitoring of LAN's, ISP networks, national infrastructures.....
- Large scale programs for Internet, from routine monitoring to ultra high resolution.
- Emergence of numerous small & large scale active probing efforts.
- Huge advances in measurement accuracy.

Renaissance in Modelling

- Return of science 101: Observation and inspiration before model.
- Return of verification: Evaluating usefulness against real data.
- New model paradigm as standard: Characteristic scale $\longrightarrow$ scale invariance.
- New model classes: Source, link, network, and closed loop.
- New problems: Active, multi-route, and high resolution measurements.

Stimulus to Theory

- Statistical estimation: For time series with infinite moments and/or scaling.
- Queueing Theory: Dealing with sub-exponential (eg LRD) input processes.
- Applied Probability: Convergence results for stochastic processes.
- Algorithms: On and Off-line synthesis methods for very long time series.
- Analysis: Network feedback, and properties of new models.

Measurement Practice

- Recognition of measurement, not just routine quantification, but discovery.
- Widespread monitoring of LAN's, ISP networks, national infrastructures.....
- Large scale programs, from routine monitoring to ultra high resolution.
- Emergence of numerous small & large scale active probing efforts.
- Huge advances in measurement accuracy.
- Router based measurements widely used.

Renaissance in Modelling

- Return of science 101: Observation and inspiration before model..
- Return of verification: Evaluating usefulness against real data.
- New model paradigm as standard: Characteristic scale $\longrightarrow$ scale invariance.
- New model classes: Source, Link, and Network.
- New problems: Active, Multi-route, and High resolution measurements.

Stimulus to Theory

- Statistical estimation: For time series with infinite moments and/or scaling.
- Queueing Theory: Dealing with sub–exponential (eg LRD) input processes.
- Applied Probability: Convergence results for stochastic processes.
- Algorithms: On & Off-line synthesis/analysis for long time series.
- Analysis: Properties of new models, especially mixed open and closed loop.

Passive and Active Measurement

Typical Passive Aims

- “At-a-point” or “Network core”.
- Backbone link utilisation, Link traffic patterns, Server workloads.
- Long term monitoring: Dimensioning, Capacity Planning, Source modelling.
- Engineering view: Network performance.

Typical Active Aims

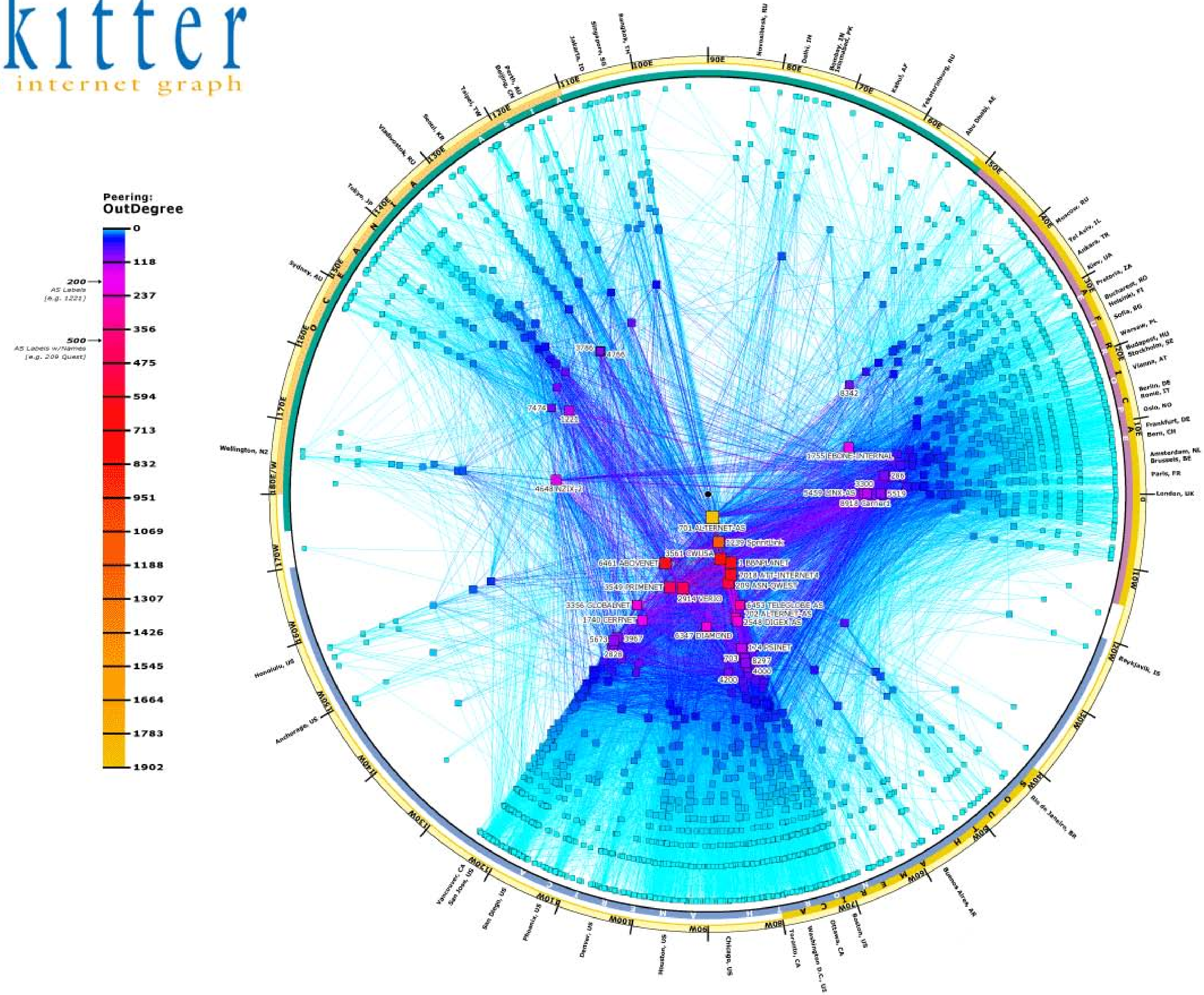
- “End-to-End” or “Network edge”.
- End-to-End Loss, Delay, Connectivity, “Discovery” and “Tomography”.
- Long and short term monitoring, “Network health” and “Route state”.
- Internet view: Application performance and robustness.

Major Measurement Programs

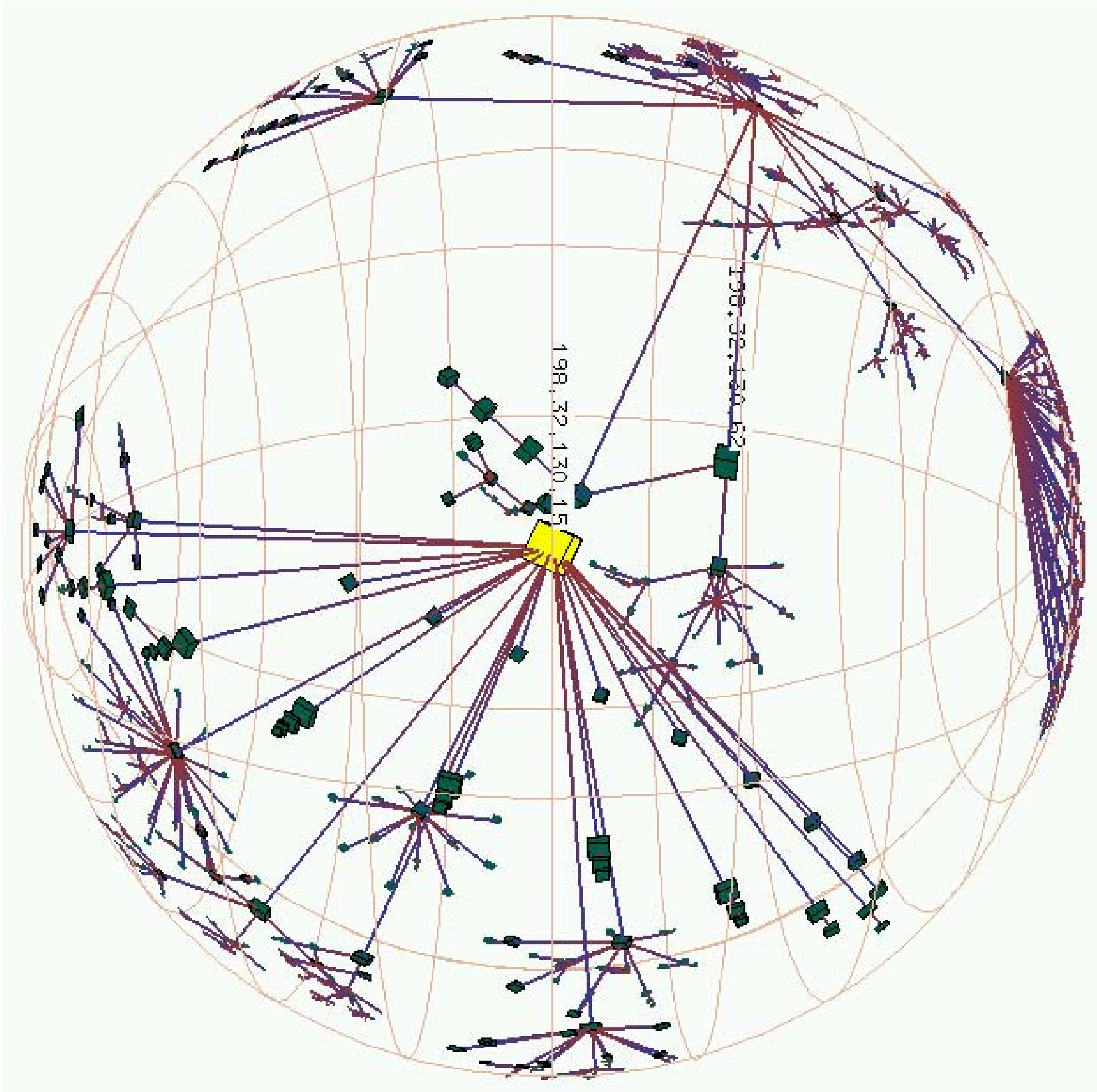
Passive Active Tools

- ● ♡ NLANR (National Laboratory for Applied Network Research).
- ♠ CAIDA (Cooperative Association for Internet Data Analysis).
- ♡ ♠ WAND (Waikato Applied Network Dynamics [DAG cards]).
- ♡ PingER (Ping End-to-end Reporting).
- ♡ Surveyor (From Advanced Networks and Services).
- ♡ ♠ RIPE–NCC (Network Coordination Centre).
- ♠ NIMI (National Internet Measurement Infrastructure).
- ● ♡ ♠ AT&T NetScope.
- ● ♠ Cicso Netflow.
- ♠ NetraMet.

skitter
AS internet graph



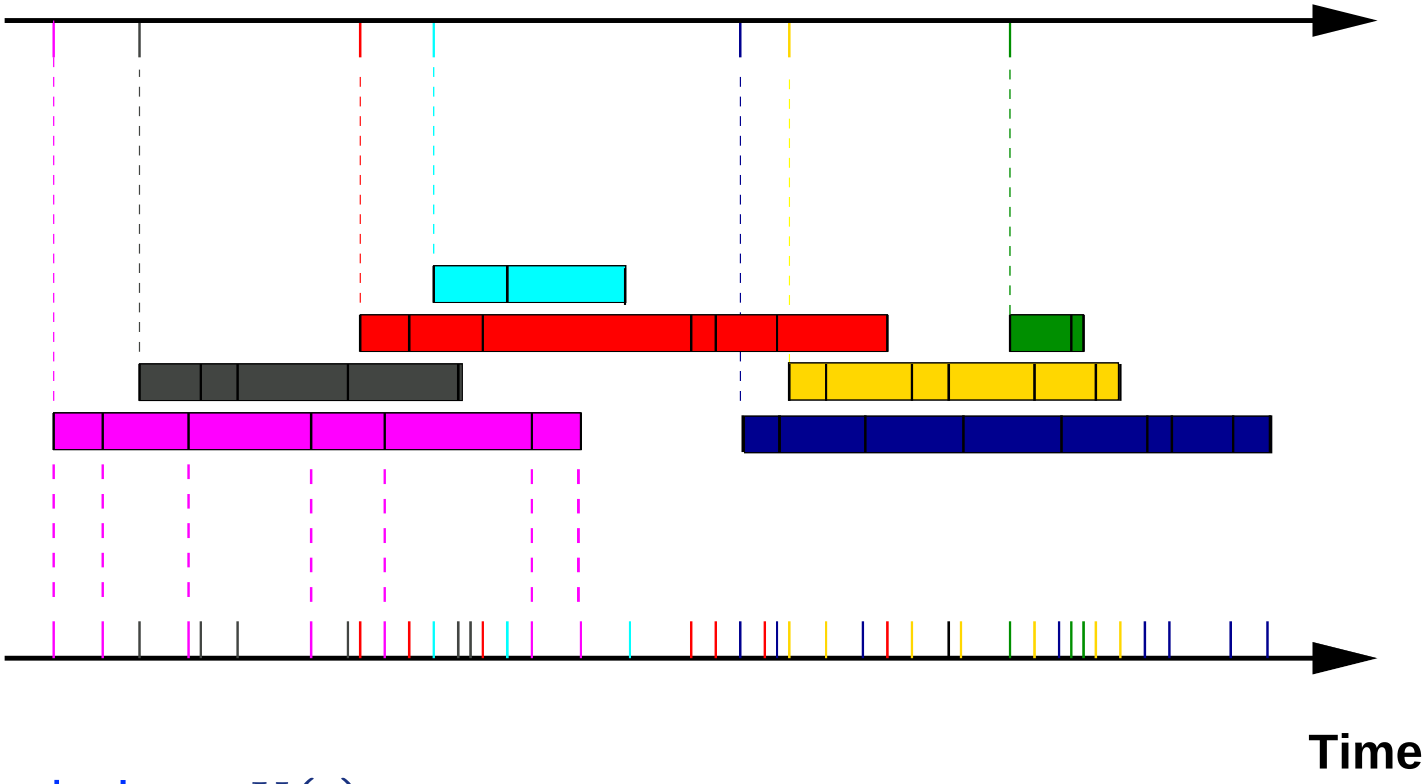
Tracing Network Nodes



Flows and Packets

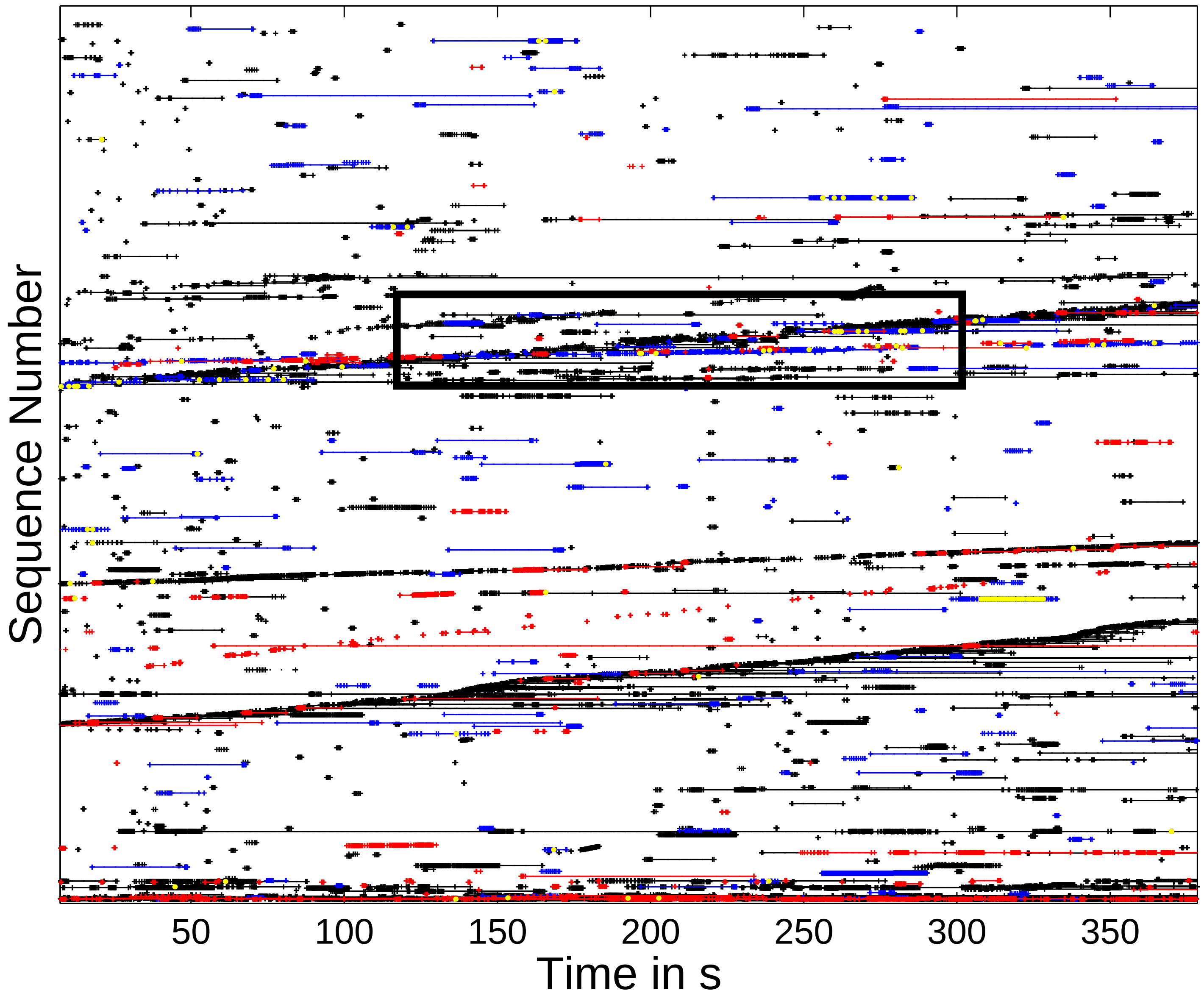
Flows are sets of packets associated to the same data transfer.

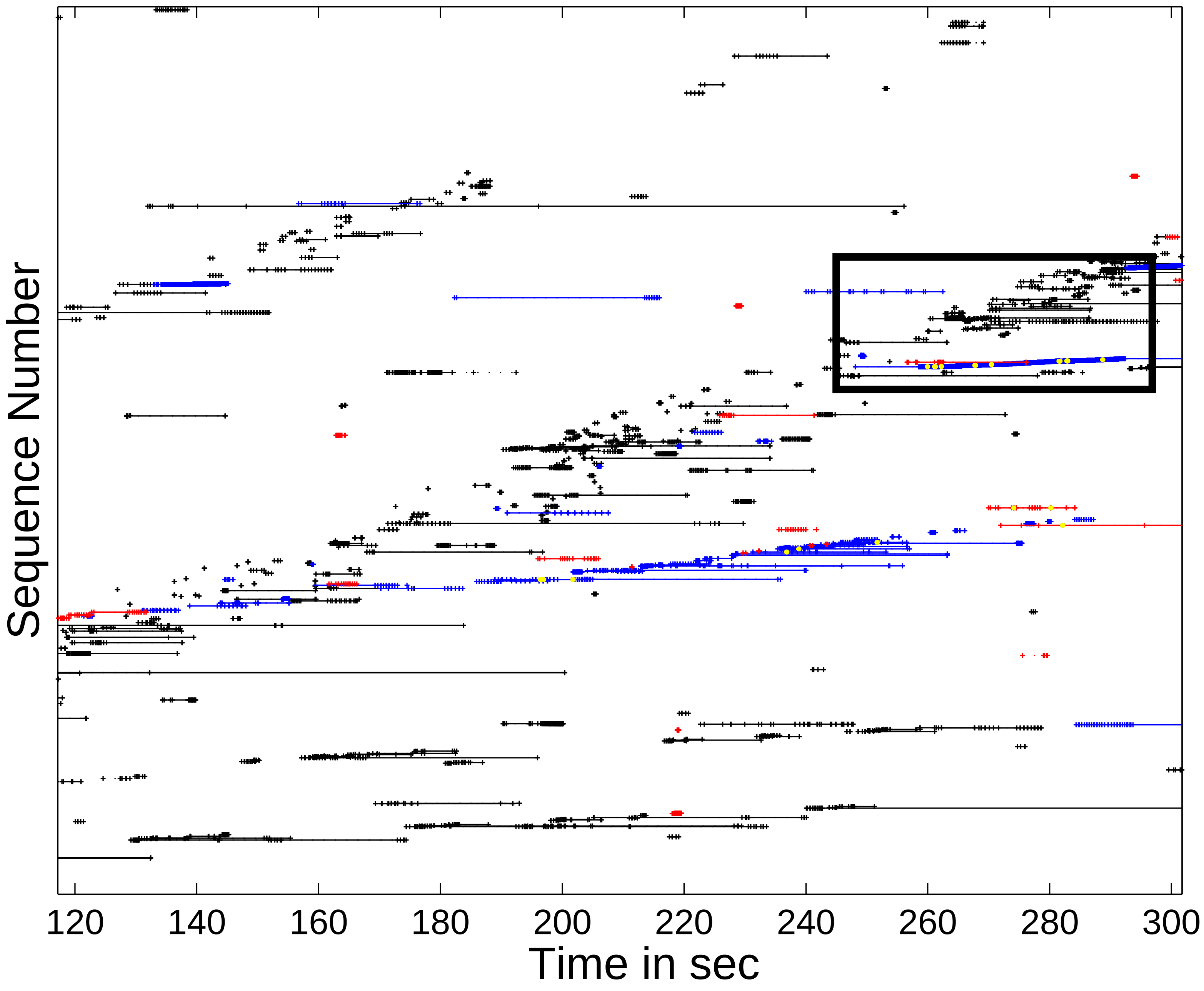
Flow arrivals: $Y(t)$



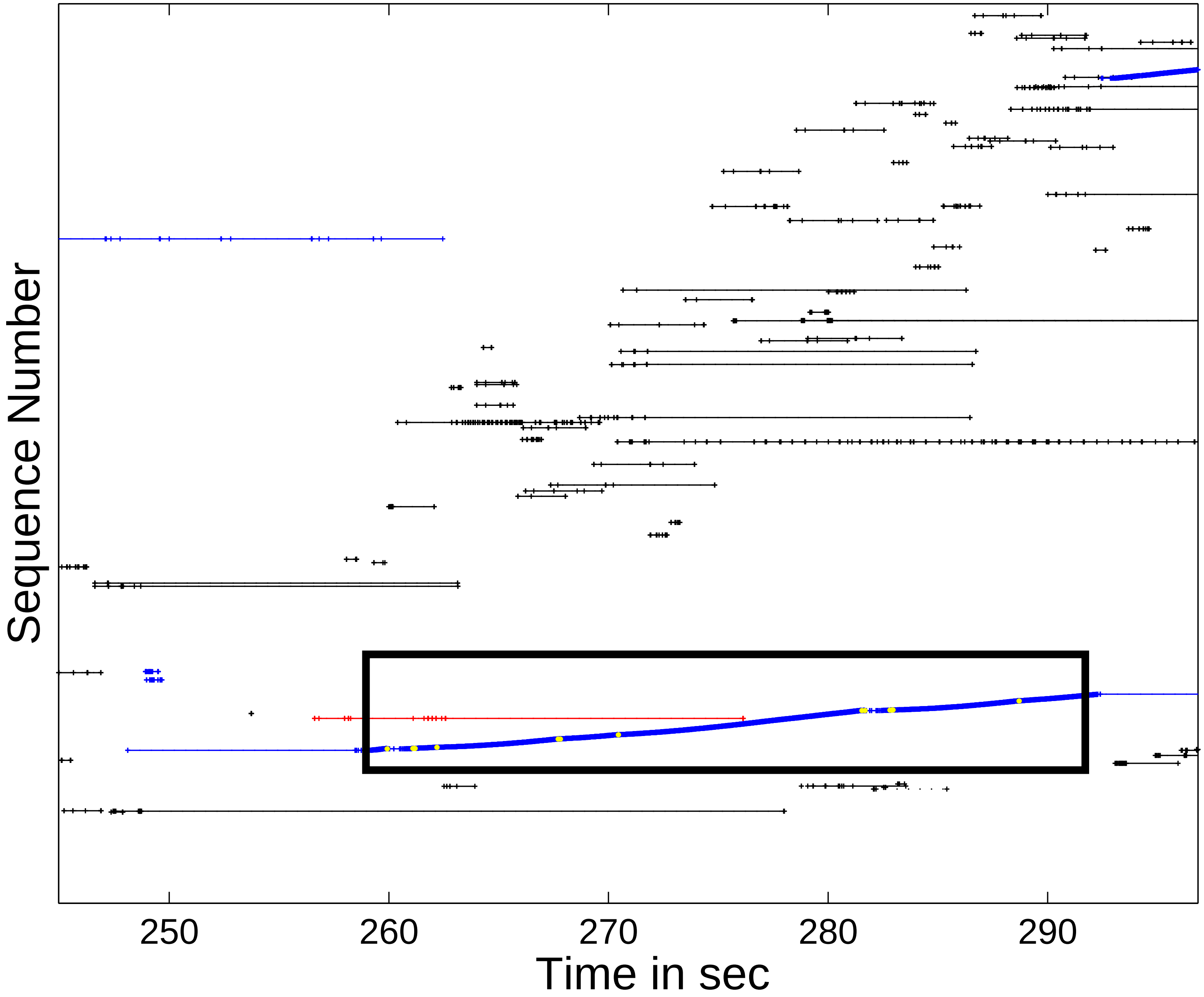
Packet arrivals: $X(t)$

Zooming on a Single Link

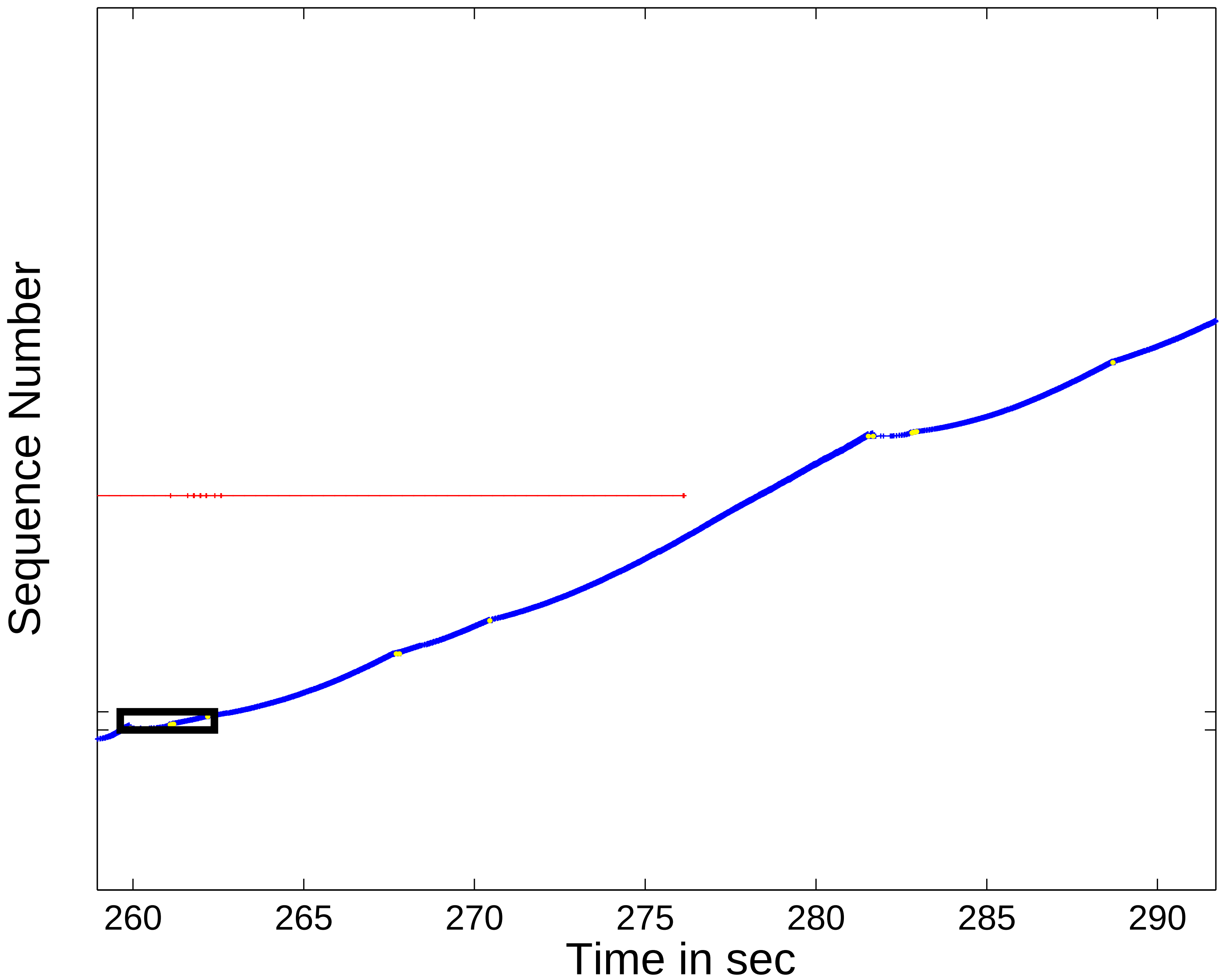


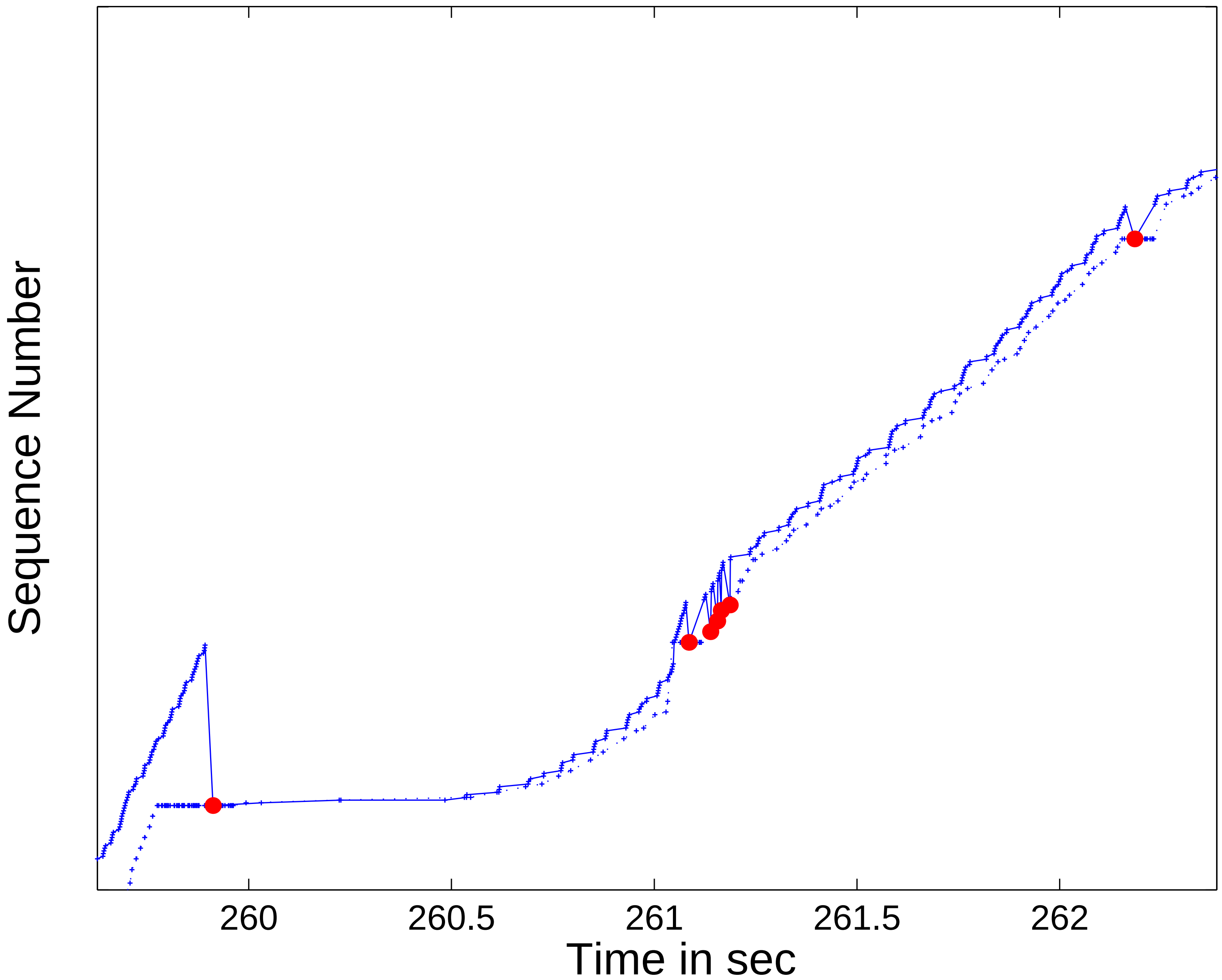


Sequence Number
120
140
160
180
200
220
240
260
280
300
Time in sec



Sequence Number
250
260
270
280
290
Time in sec


Sequence Number
260
265
270
275
280
285
290
Time in sec



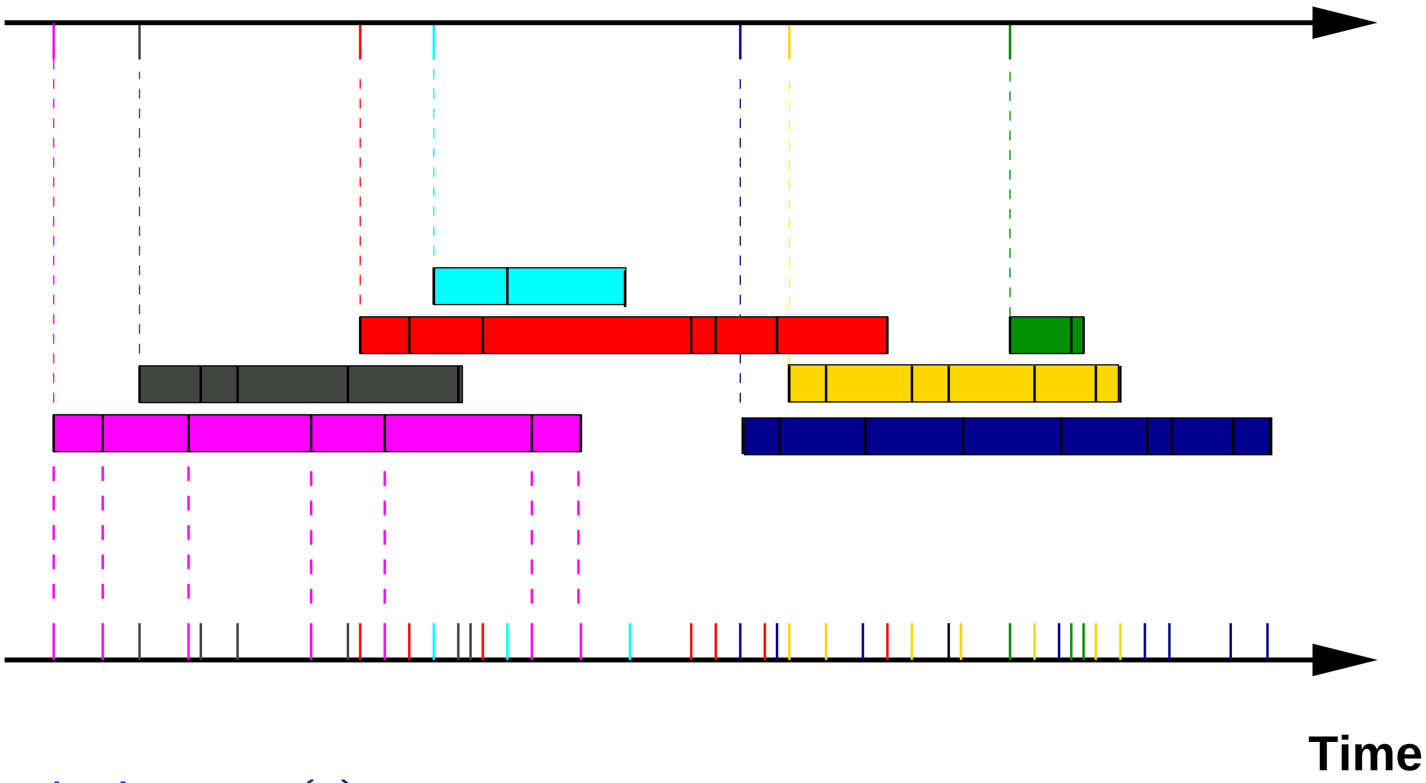
Sequence Number
260
260.5
261
261.5
262
Time in sec

What are we Measuring?

- *Internet Protocol* (IP) packets, the unit of transport across networks.
- Data split into packets, with: header, payload.
- Payload carries higher layer protocols: TCP, UDP, ICMP.
- Protocols support services & applications, more protocols:
  - TCP: HTTP, FTP, SNMP, ... (for reliable data)
  - UDP: VoIP, DNS, NTP,... (for real time)
- For each packet:
  - Could filter based on criteria (address, type, size, ...)
  - Could capture all or part (e.g. just the header).
  - Must timestamp.
- Key concept, a flow (collection) of packets.

Two Point Processes from Traffic

Flow arrivals: $Y(t)$



Packet arrivals: $X(t)$

Most common time series extracted are packets or bytes per bin

Seeing Packet Traffic

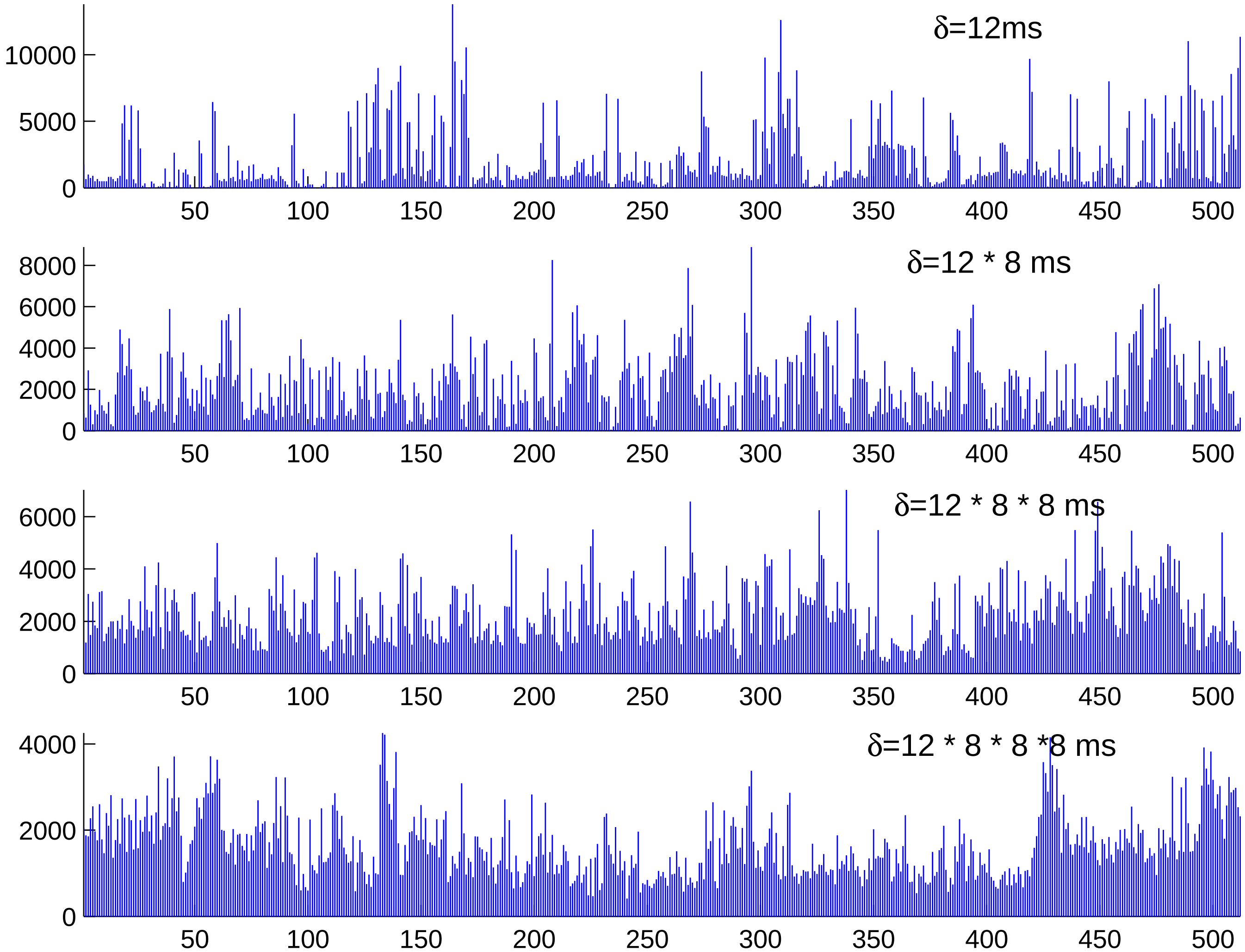
- 1991: ISDN (Hellstern, Wirth, Yan, Hoeflin) – infinite moments.
- 1991: Ethernet (Leyland, Wilson) – bursts at all time scales.
- 1993: Ethernet (Erramilli, Willinger) – fractal properties.

Believing Packet Traffic

- 1993: Ethernet (Leyland, Taqqu, Willinger, Wilson) – Self-Similar Traffic.
- 1994: CCSN/SS7 (Duffy, McIntosh, Rosenstein, Willinger) – Self-Similar.
- 1994: Internet (Paxson, Floyd) – Failure of Poisson modelling.
- 1994 $\longrightarrow \cdots$ LAN's across the world see – Self-Similarity.
- 1994: Web docs (Cunha, Bestavros, Crovella) – Heavy tailed file sizes.
- 1995: Video (Beran, Sherman, Taqqu, Willinger) – roughly Self-Similar.

The Self-Similarity of Ethernet Traffic

The reference Bellcore trace, ‘pAug’, is close to *Fractional Gaussian Noise*.



Measuring the Exponent using Wavelets

Wavelet coefficients of our traffic process:

$$d_X(j,k) = \langle X, \psi_{j,k} \rangle.$$

Spectral definition of LRD is

$$\Gamma_X(\nu) \sim c_f |\nu|^{-\alpha}, \; |\nu| \to 0, \text{ with } \alpha \in (0,1).$$

In this case, provided $N \geq 1$ vanishing moments:

$$\mathbf{E}|d_X(j,k)|^2 \sim c_f C(\alpha)\, 2^{\alpha j}, \; j \to +\infty$$

Estimating the LHS from data using

$$S_2(j) = \frac{1}{n_j} \sum_k |d_X(j,k)|^2,$$

where n_j is the number of $d_X(j,k)$ available at octave j (scale $a = 2^j$),

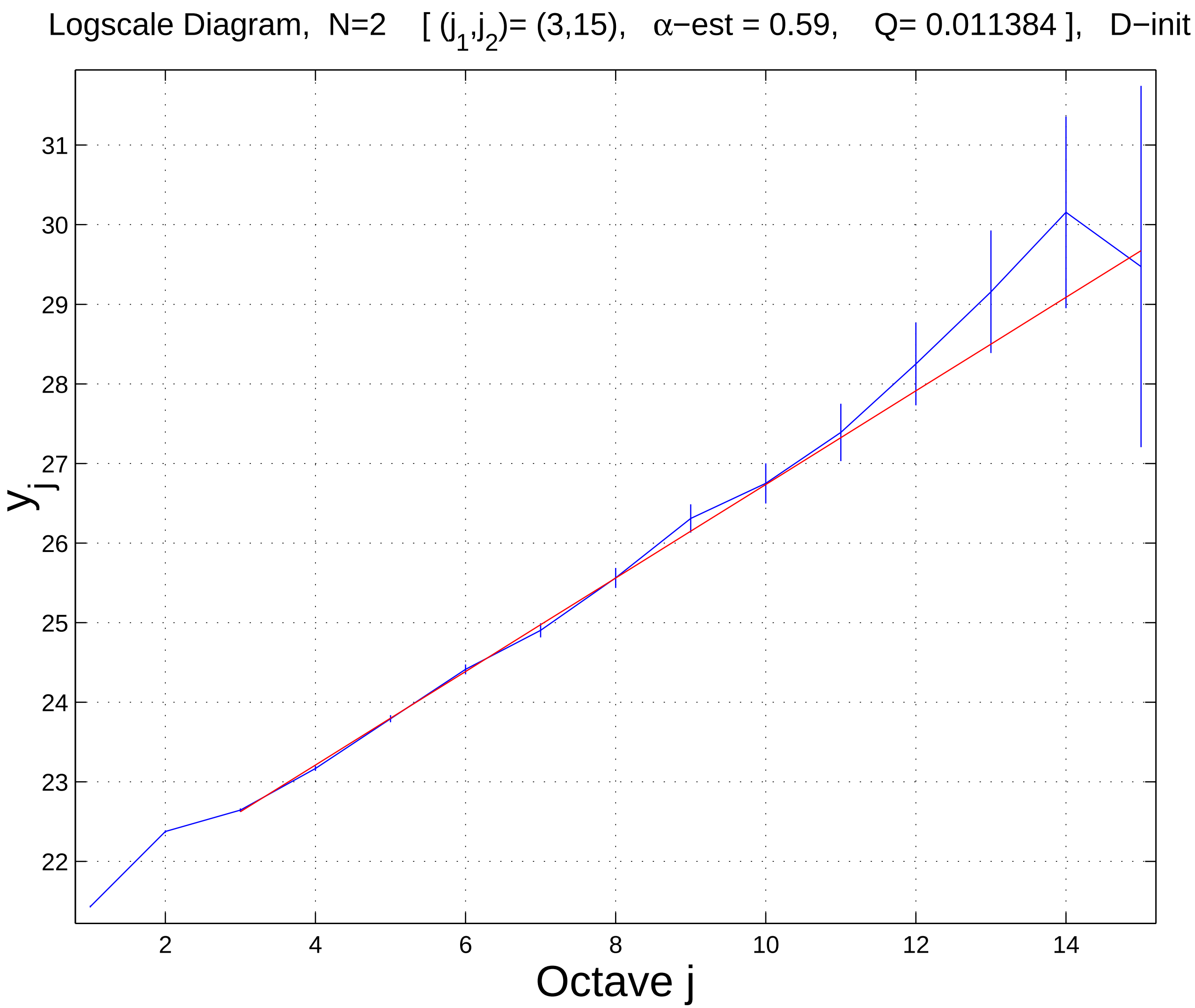
We use '*Logscale Diagram*' to refer to the log-log plot:

$$\text{LD}: \quad \log_2 S_2(j) \text{ vs } j,$$

in which straight lines are evidence for scaling, slope$= \alpha$.

Analysis of Trace ‘pAug’

ETHERNET: bytes per 12ms bin.



What does it mean?

- No characteristic time-scale controlling the dynamics / statistics.
- Statistically, All scales (in a range) are equivalent, under a renormalisation.
- Radical temporal burstiness: no natural burst scale.
- Key parameters no longer special scales, but relations across scales.
- Absolute quantities $\longrightarrow$ scale dependent quantities.
- Affects system description, behaviour, measurement ....

The Self-Similar Traffic Model

Fractional Gaussian Noise (fGn) and Fractional Brownian Motion (fBm)

The unique Gaussian processes which are

$$\begin{aligned} \text{Stationary; Cov } [X_H](k) &= \frac{1}{2}\Big[(k-1)^{2H} + 2k^{2H} + (k+1)^{2H}\Big] \\ \text{Stationary Increments; Var } [Z_H(k)] &= k^{2H} \end{aligned}$$

Corresponding traffic models:

$$\begin{aligned} \text{Rate: } R(t) &= \mu + \sigma X_H(t) \\ \text{Total Traffic: } W(t) &= \mu t + \sigma Z_H(t) \end{aligned}$$

where $Z_H(t) = \sum_{i=1}^{t} X_H(i), \quad W(t) = \sum_{i=1}^{t} R(i).$

The Long–Range Dependent (LRD) Traffic Models

LRD definition: a slowly decaying covariance

$$\Gamma_X(k) \sim c_r k^{-\beta}, \quad 0 < \beta < 1,$$

where $(\beta = 2 - 2H)$.

Corresponding traffic rate model:

$$R(t) \;=\; \mu + \sigma X_{\beta,c_r,k^*}(t).$$

LRD more general than H-SS:

- Second order description only, not necessarily Gaussian!
- Has been called second order asymptotically self-similar (but careful!)
- Heavy tail 'begins' only after some cutoff scale k^*.
- Tail may be 'thin', low mass (small c_r), independent of variance!
- Small scale structure unspecified.
- At a minimum, three correlation parameters, not just H.

NATIONAL ICT AUSTRALIA LIMITED

Non-Gaussian LRD – the On/Off Source

- Alternating renewal process: $\{A_i\}$ i.i.d. $\{B_i\}$ i.i.d.
- LRD if A or B heavy tailed:
  - If $\Pr(B > x) \sim c\, x^{-\alpha}$, $\beta_{\mathrm{LRD}} = 3 - \alpha$, $c_r = \frac{c(1-\lambda)^3}{(\alpha-1)\mathbf{E}[A]}$.
  - Efficient generation ($O(1)$ computation and state)

Corresponding traffic rate models:

- as active/silence sources.
- as building blocks for a compound source, eg.
  - $N \uparrow$, $p = \Pr(\mathrm{On}) \to 0$, $\lambda =$ const, $h =$const: $\longrightarrow$ M/G/∞
  - $N \uparrow$, $p =$const, $\lambda =$ const, $h \to 0$: $\longrightarrow$ fGn

Impacts on Traffic Modelling

Statistics

- Difficult estimation of almost everything.
- Difficult model choice: Non-stationarity confusions and others.
- Large confidence intervals (need more data).
- Wavelets can help kill LRD, mitigate NS, in many cases.
  Complexity only $O(n)$ and can be done on-line.

Queueing

- Larger buffer sizes, lower utilisation at given QoS.
- 'Buffer insensitivity' (bigger won't help much).
- Impact depends a lot on the amplitude burstiness:
  - fBm storage (Norros) – Weibullian tail.
  - large peak On/Off superpositions – Regularly varying tails, infinite mean (Cohen; Boxma; Whitt).
- Service discipline is important – can reinstate finite mean (Boxma et al.)
- Critical time scales can exist, but value depends on full structure.

Lessons for Network Performance

- Heavy tails can be thinned, but not cut off.
- Smoothing the beast will not tame it.
- Statistical Multiplexing works as ever – recommended.
- Processor Sharing helps the individual, but not the society.
- Hi Peak rates can kill.
- Power-laws are very persistent, if you can't kill it:

  Make sure it is irrelevant, or

  Make it work for you, or

  Understand it, calm it down, then live with it.
- LRD not everything, obviously: variance, *marginal*.

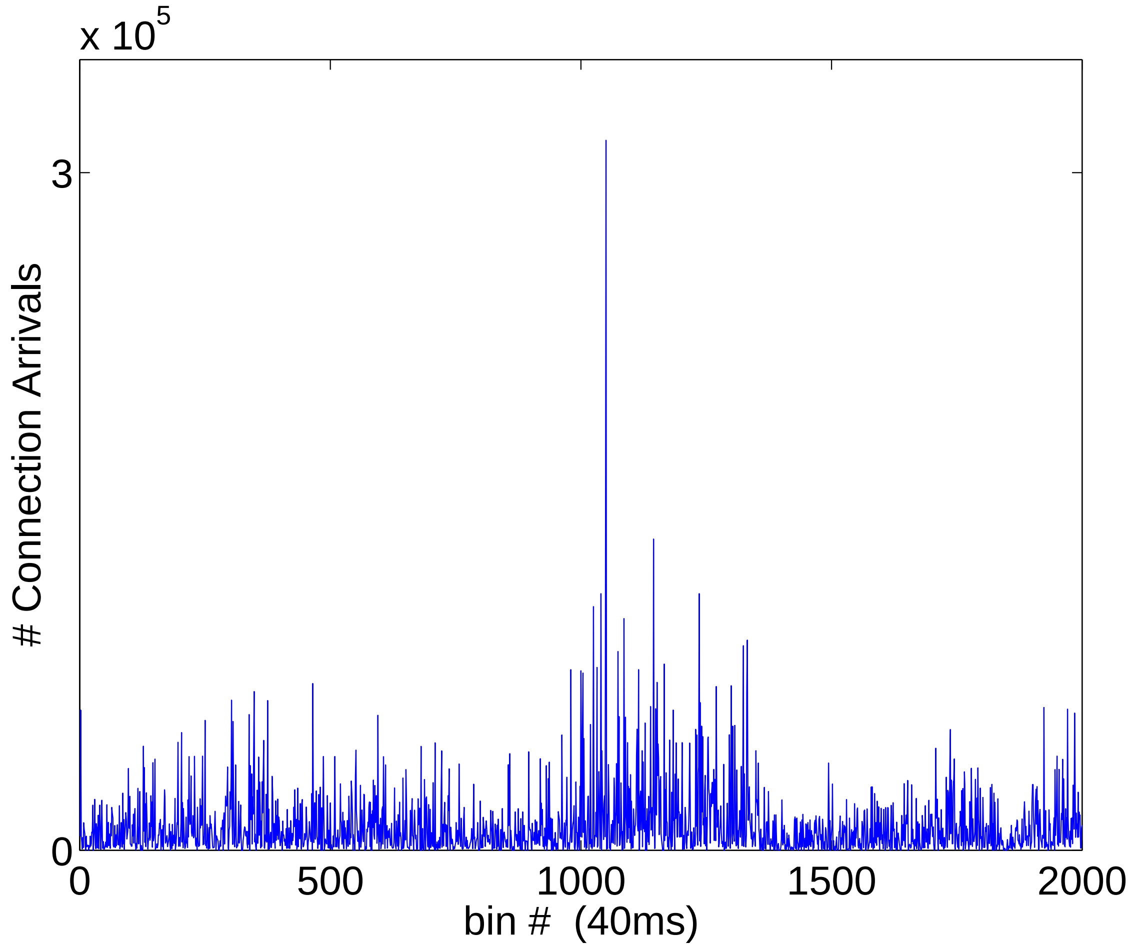
NATIONAL ICT AUSTRALIA LIMITED

BUT: Internet, TCP, Small Scales, and Non-Gaussianity

The ‘Self-Similarity’ model has limits

- Only really asymptotic: true for ‘large scales’, $\approx >$ 1 second.
- Really requires Gaussianity – often far from the case!
- Range of scales where valid may not be crucial for engineering purposes.
- At small scales, feedback can mold the beast.

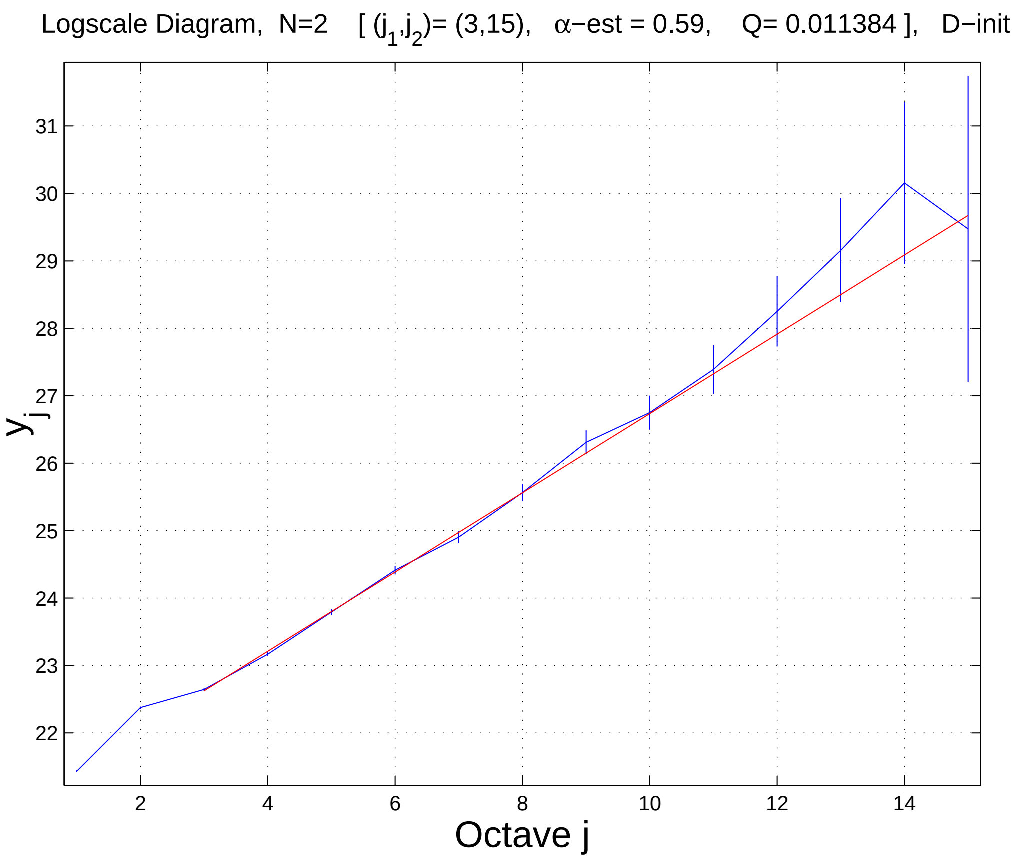
High Amplitude Burstiness at Small Scale



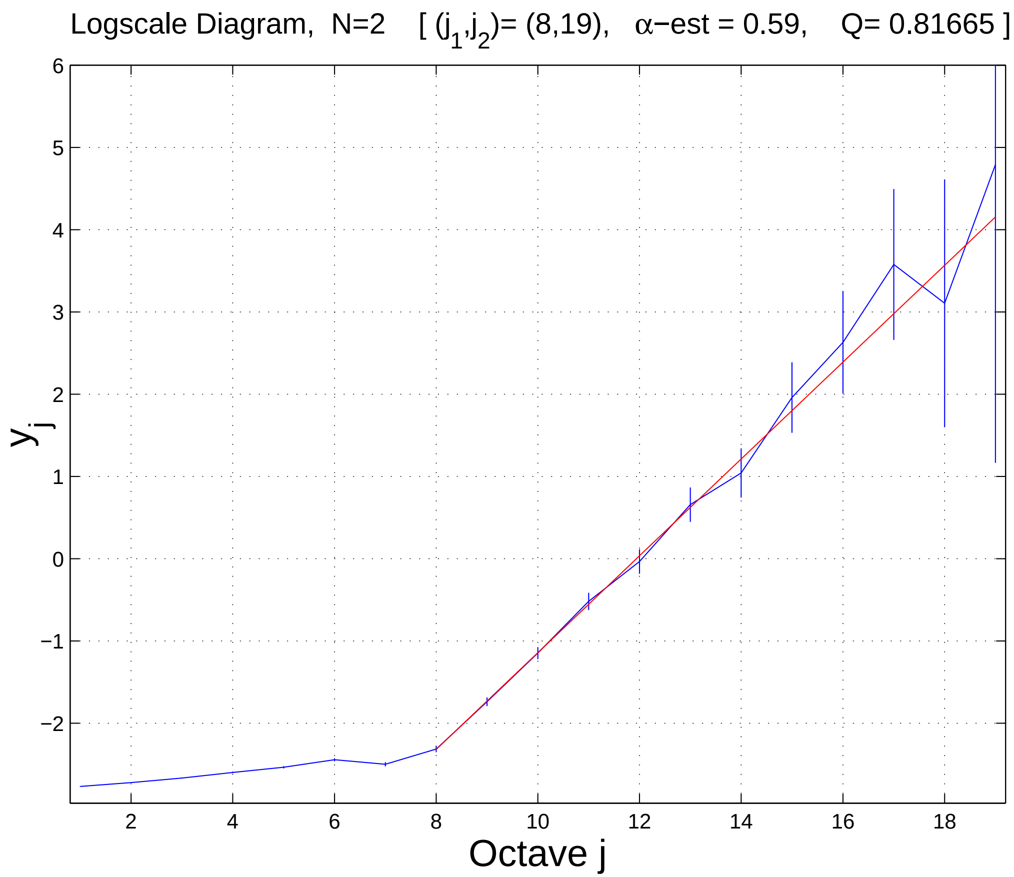
So small scales are much harder ...

Examples of Scaling in Traffic: 2nd Order Wavelet Analysis

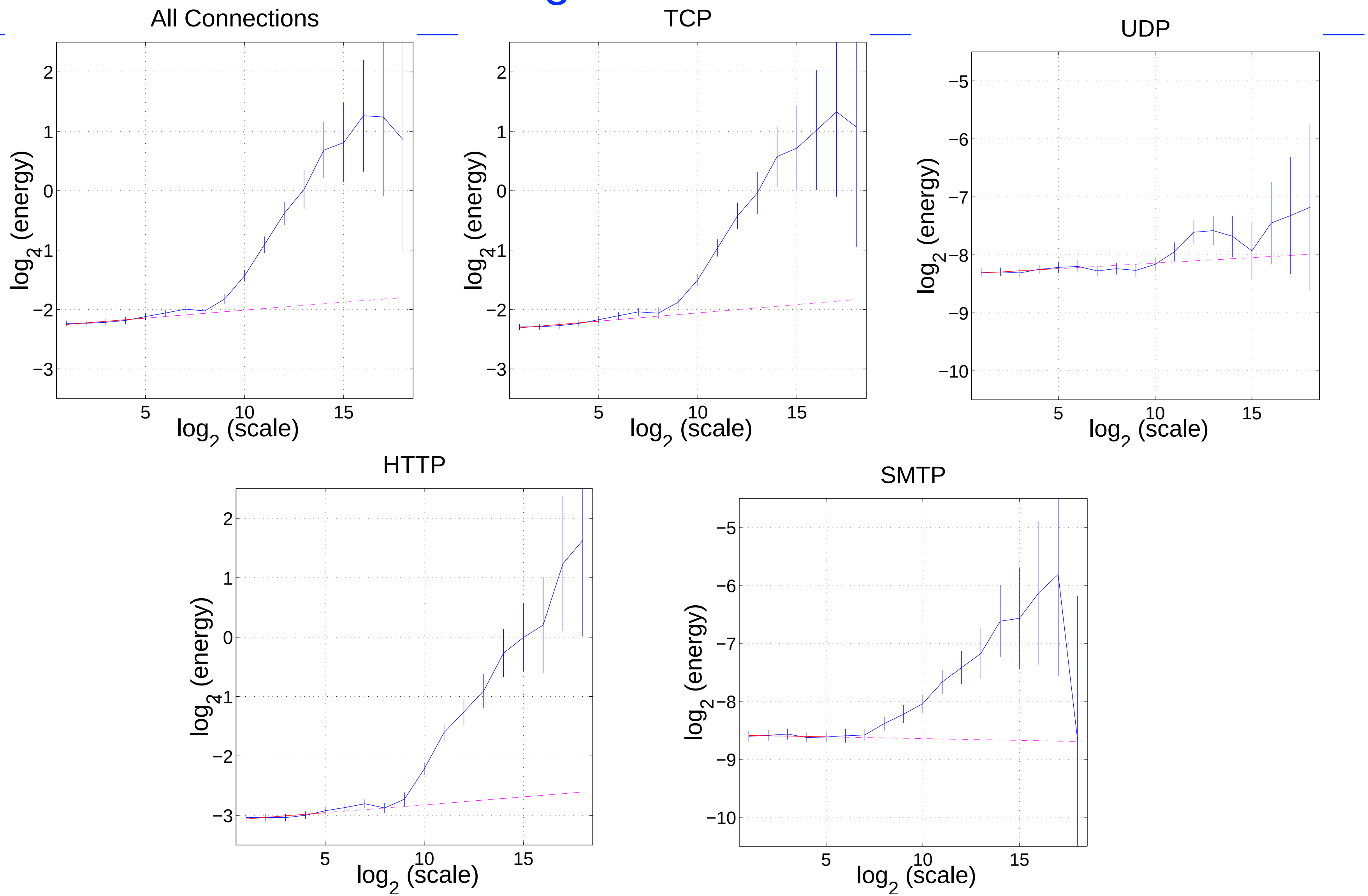
ETHERNET: bytes per 12ms bin.



INTERNET: new connections per 10ms bin.



Biscaling in TCP Arrivals



Collaboration with Patrice Abry and Nicolas Hohn

Evidence for Multifractality

Wavelet qth order moments: $\mathbf{E}|d_X(j,k)|^q \sim C\,2^{\alpha_q j},\ j \to 0.$

Estimating the LHS from data using

$$S_q(j) = \frac{1}{n_j}\sum_k |d_X(j,k)|^q,$$

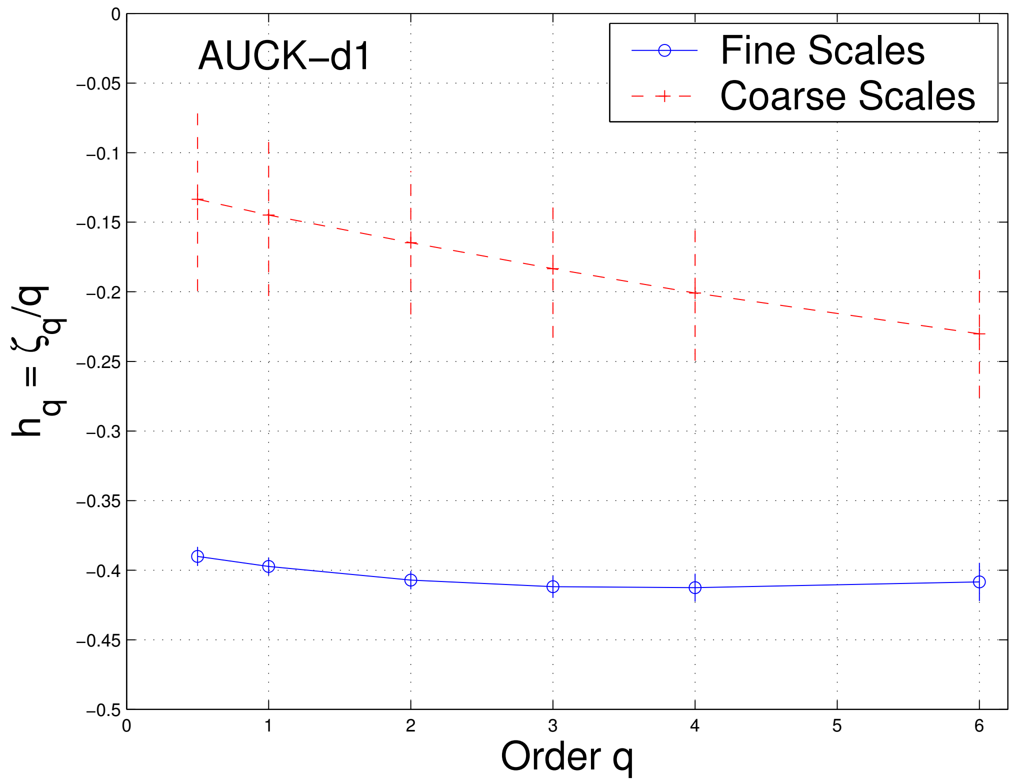
and measure the slopes $\widehat{\alpha}_q = \zeta(q) + q/2$ in a log-log plot (the ‘MD‘).

Instead of testing for linearity of $\zeta(q)$ *vs* q, look for $\zeta(q)/q$ *vs* q constant.

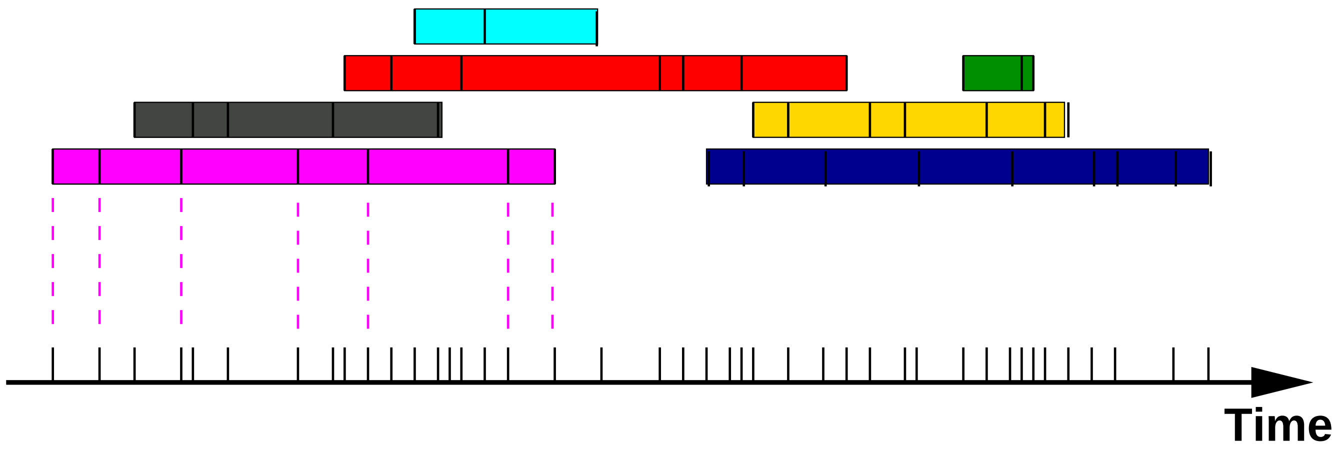
A Short History of Multifractal Traffic Modelling

- 1997: Ethernet (Taqqu, Teverovsky, Willinger) – not MF
  time domain ‖ discrete packet and byte counts ‖ $\delta = 10$ms.
- 1997: TCP, LAN gateway (Riedi, Véhel) – MF at 'high freq'
  increments ‖ packet sizes, iat's, packet and byte counts ‖ $\delta = 150$ms.
- 1998: Ethernet (Abry, Veitch) – not MF
  wavelet distributions ‖ continuous time byte counts.
- 1997-8: TCP, LAN gateway (Feldmann, Gilbert, Willinger, Kurtz) – MF small, Mono large.
  wavelet domain; discrete packet counts ‖ $\delta = 10$ms ‖ large & small regimes
- 1998: TCP, WAN (Feldmann, Gilbert, Willinger) – MF at small, Mono large
  wavelet domain ‖ discrete packet counts ‖ $\delta = 10$ms.
- 1999: ISP and simulated (Feldmann, Gilbert, Huang, Willinger) – MF at small, but dirty
  wavelet domain ‖ discrete packet counts ‖ $\delta = 10$ms.
- 1999-2000: TCP, WAN (Veitch, Abry, Flandrin, Chainais) – IDC holds over large and small.
  wavelet distributions ‖ TCP connection counts ‖ $\delta = 10$ms ‖ CI's used.
- 2001: TCP, WAN (Roux, Veitch, Abry, Huang, Flandrin, Micheel) – IDC, but $\approx$mono.
  wavelet distributions ‖ TCP connection counts, packet iar's and counts, byte counts $\delta = 10$ms ‖ CI's used
- 2003: TCP, WAN, high rate (Zhang, Ribeiro, Moon, Diot) – Mono everywhere
  wavelet domain ‖ discrete byte counts ‖ $\delta = 10$ms. ‖ CI's used
- 2003: TCP, WAN, high rate (Hohn, Veitch, Abry) – Empirical scaling misleading?
  wavelet domain ‖ continuous packet counts ‖ $\delta = 5\mu s \rightarrow 5ms$ ‖ CI's used ‖ $q = 2$

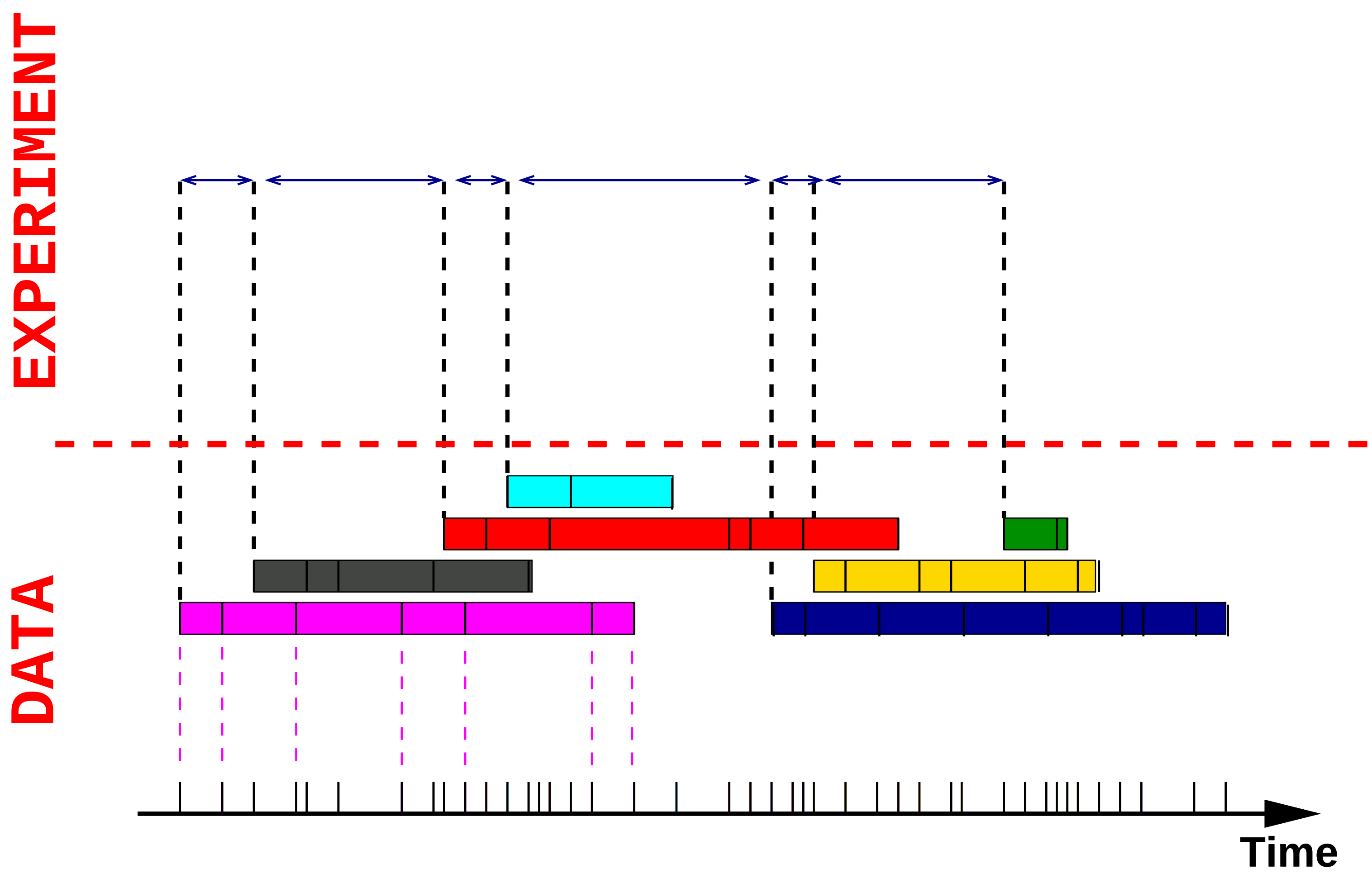
Similar Data, A Less Flattering View



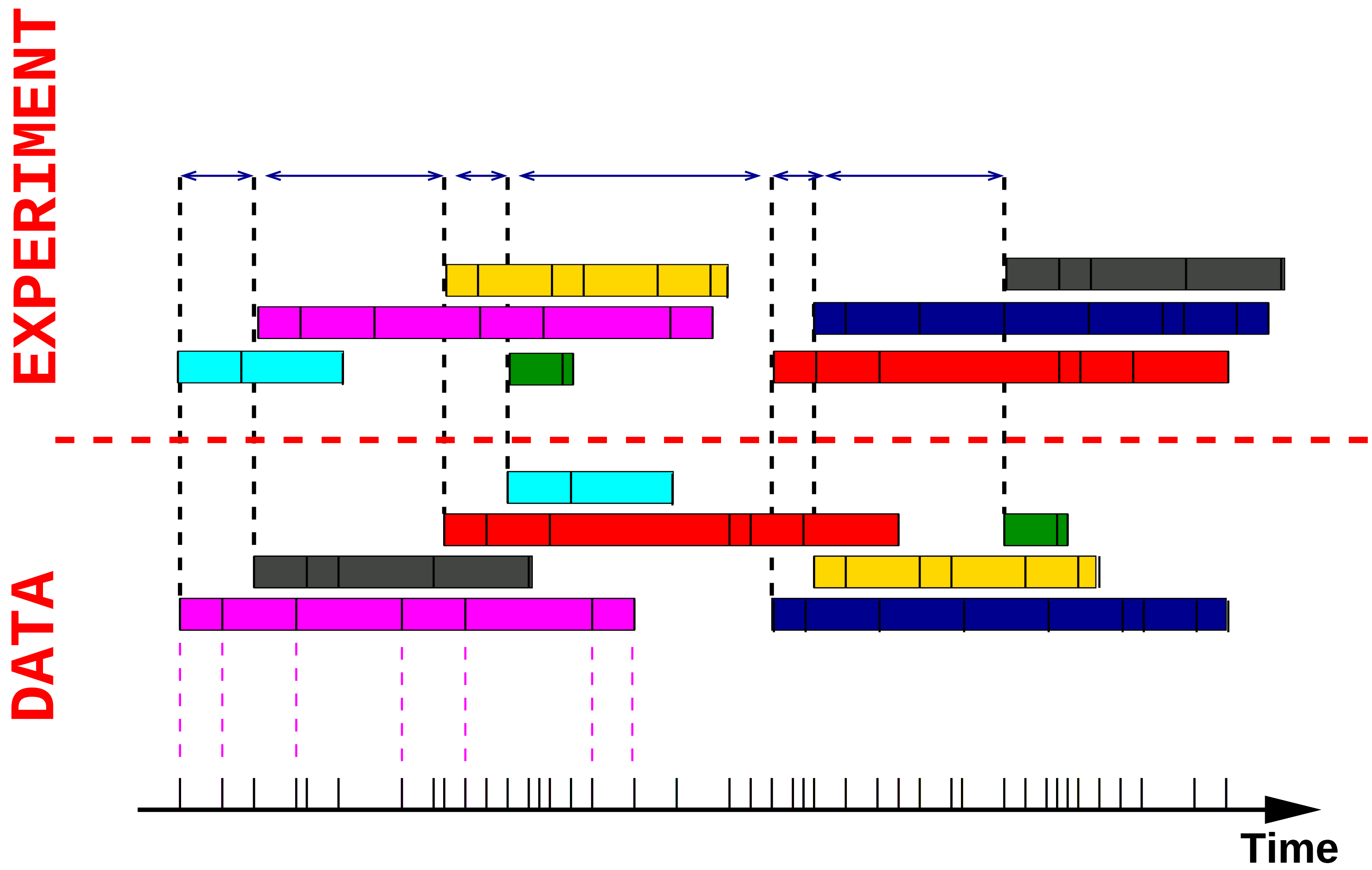
Semi-Experiments on Packet Arrivals



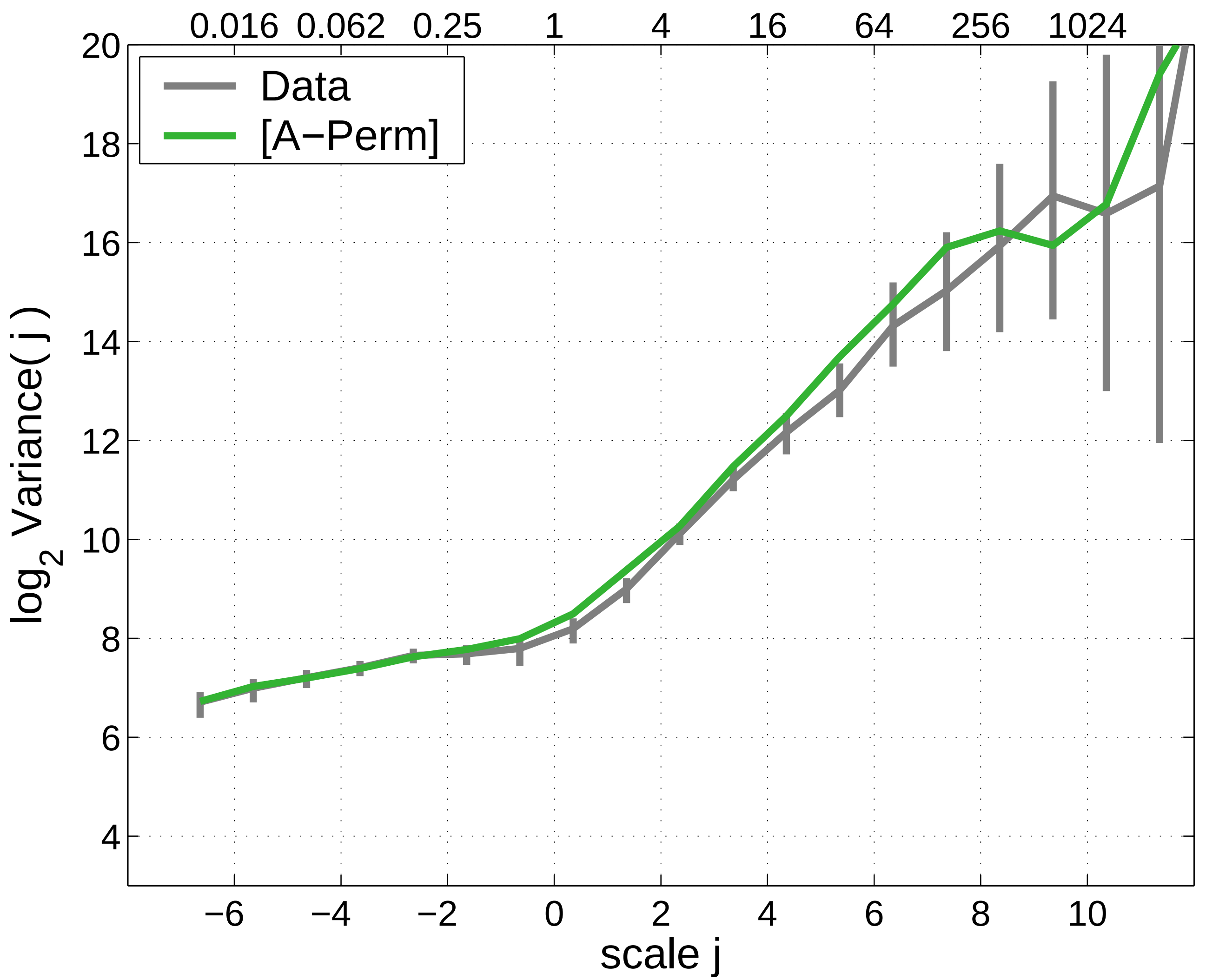
Original TCP Data



Permutation of Flows [A-Perm]

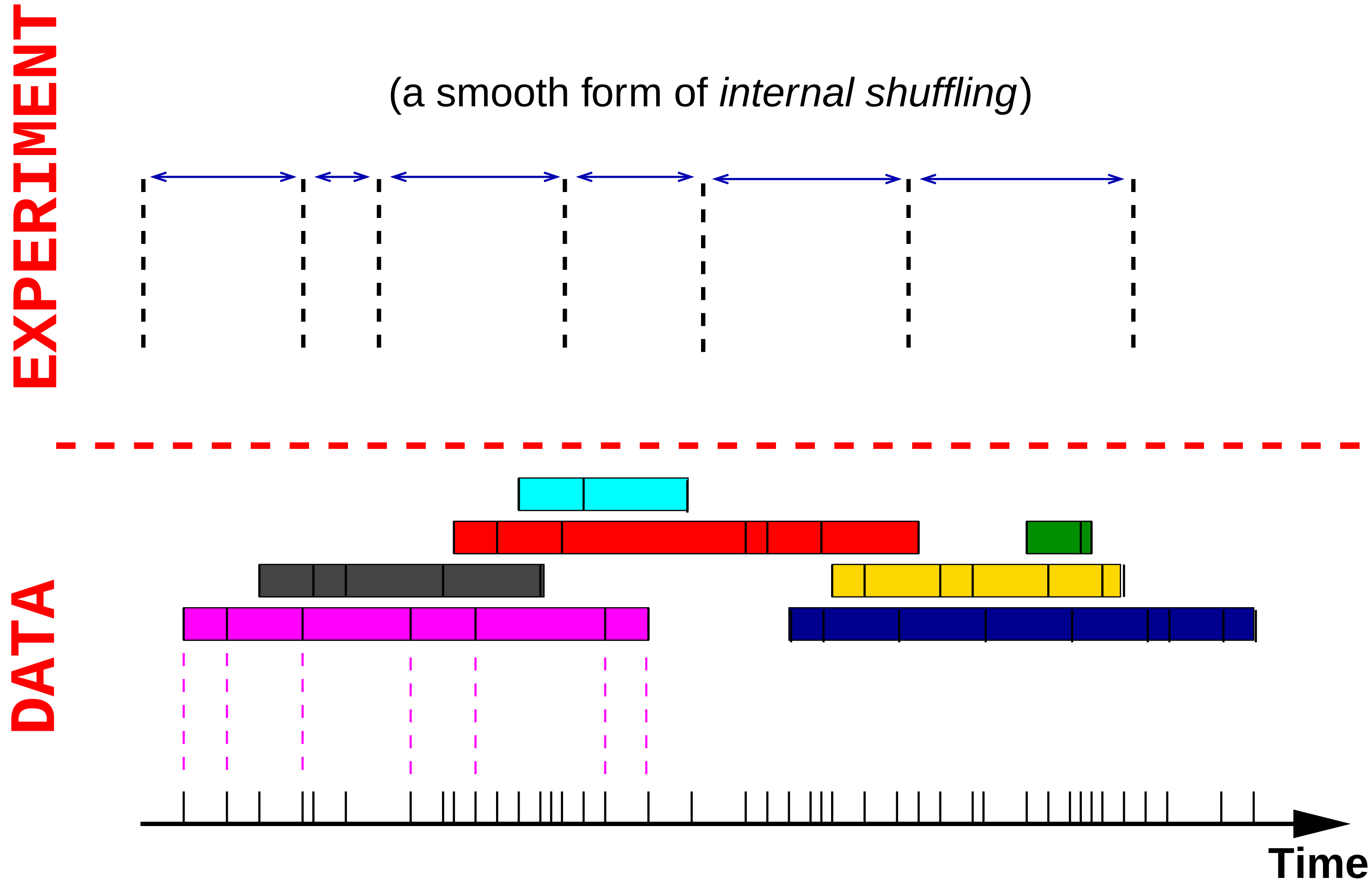


Permutation of Flows [A-Perm]

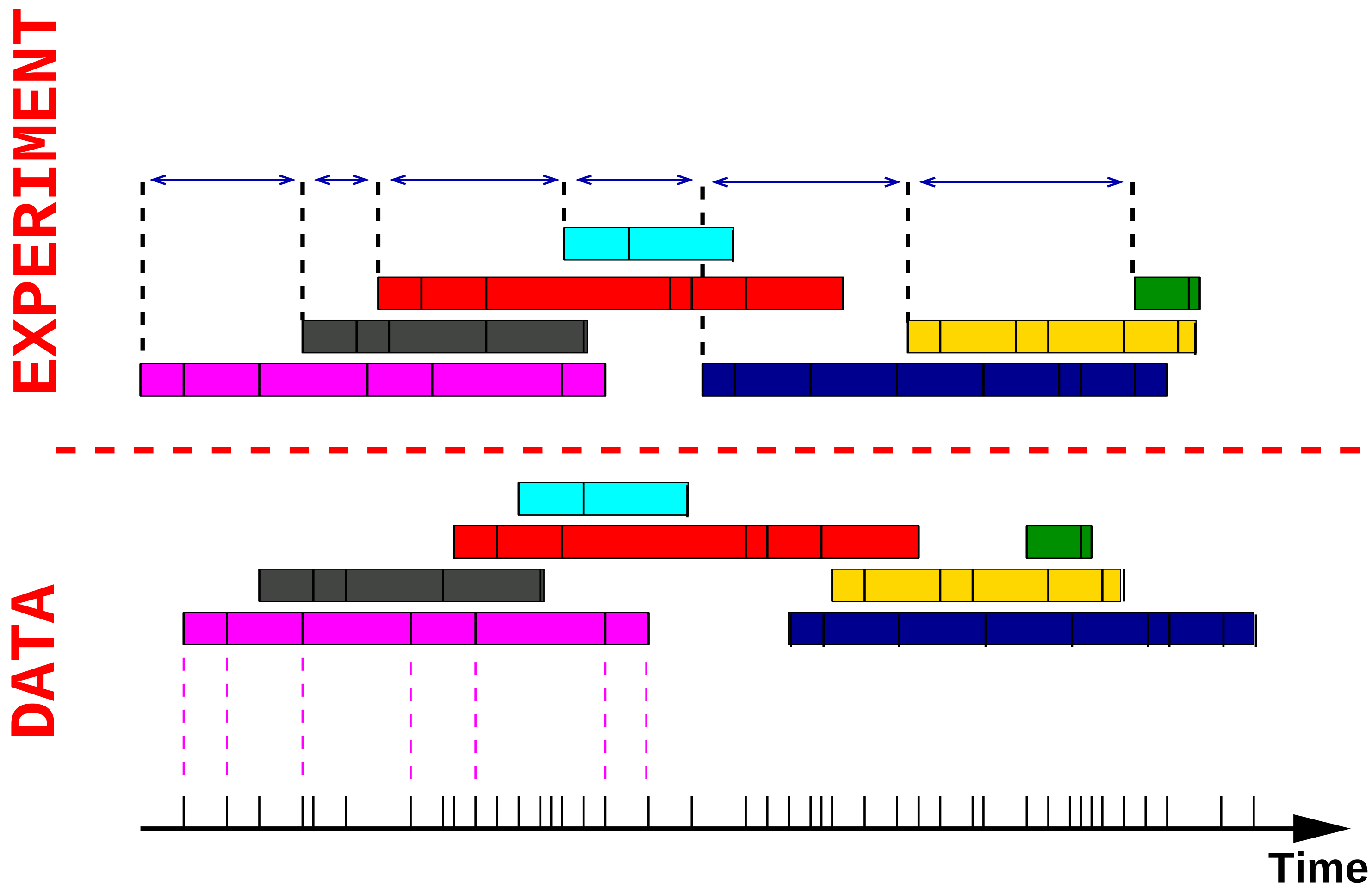


Original Order Poisson Arrivals [A-Pord]

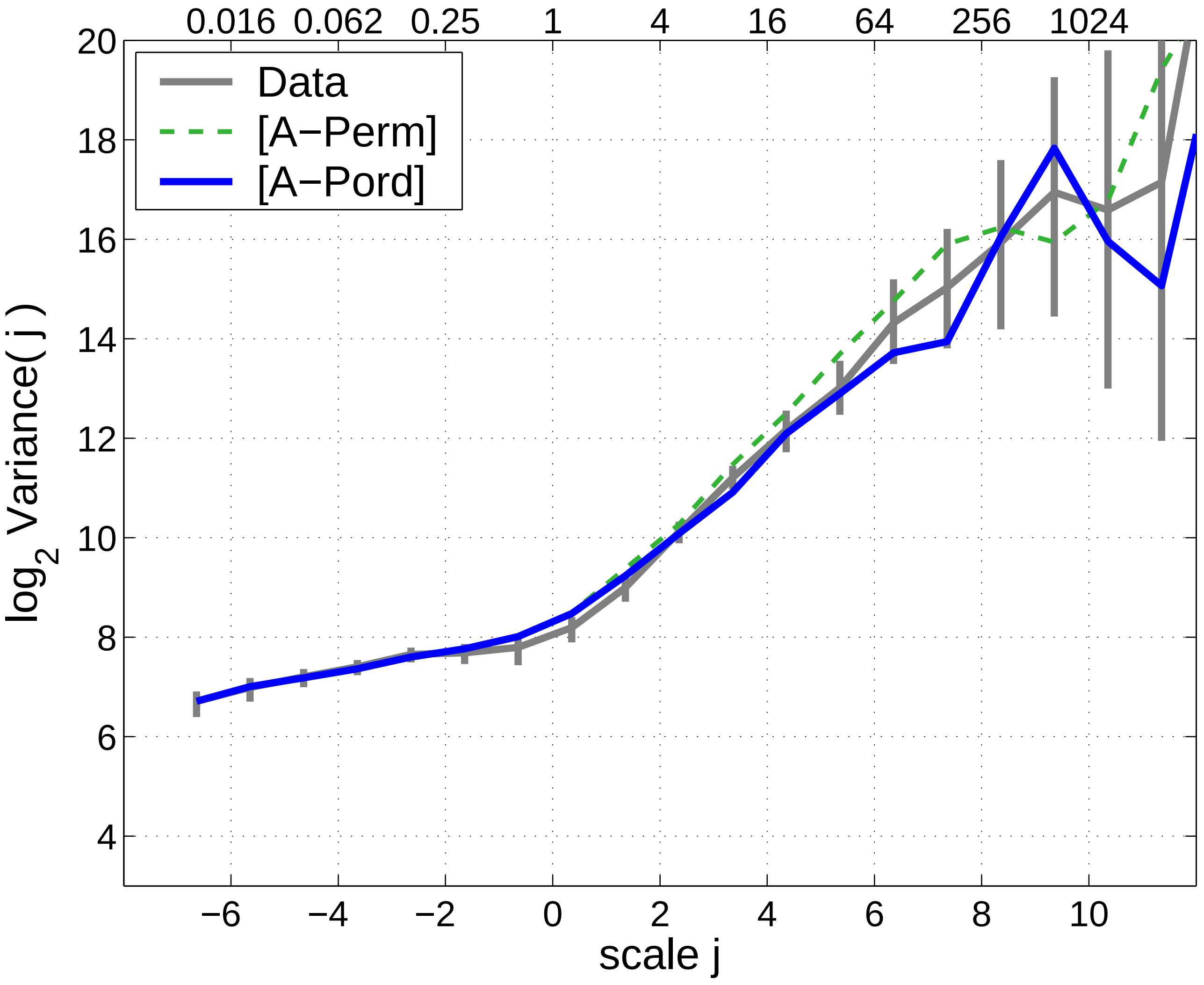
(a smooth form of *internal shuffling*)



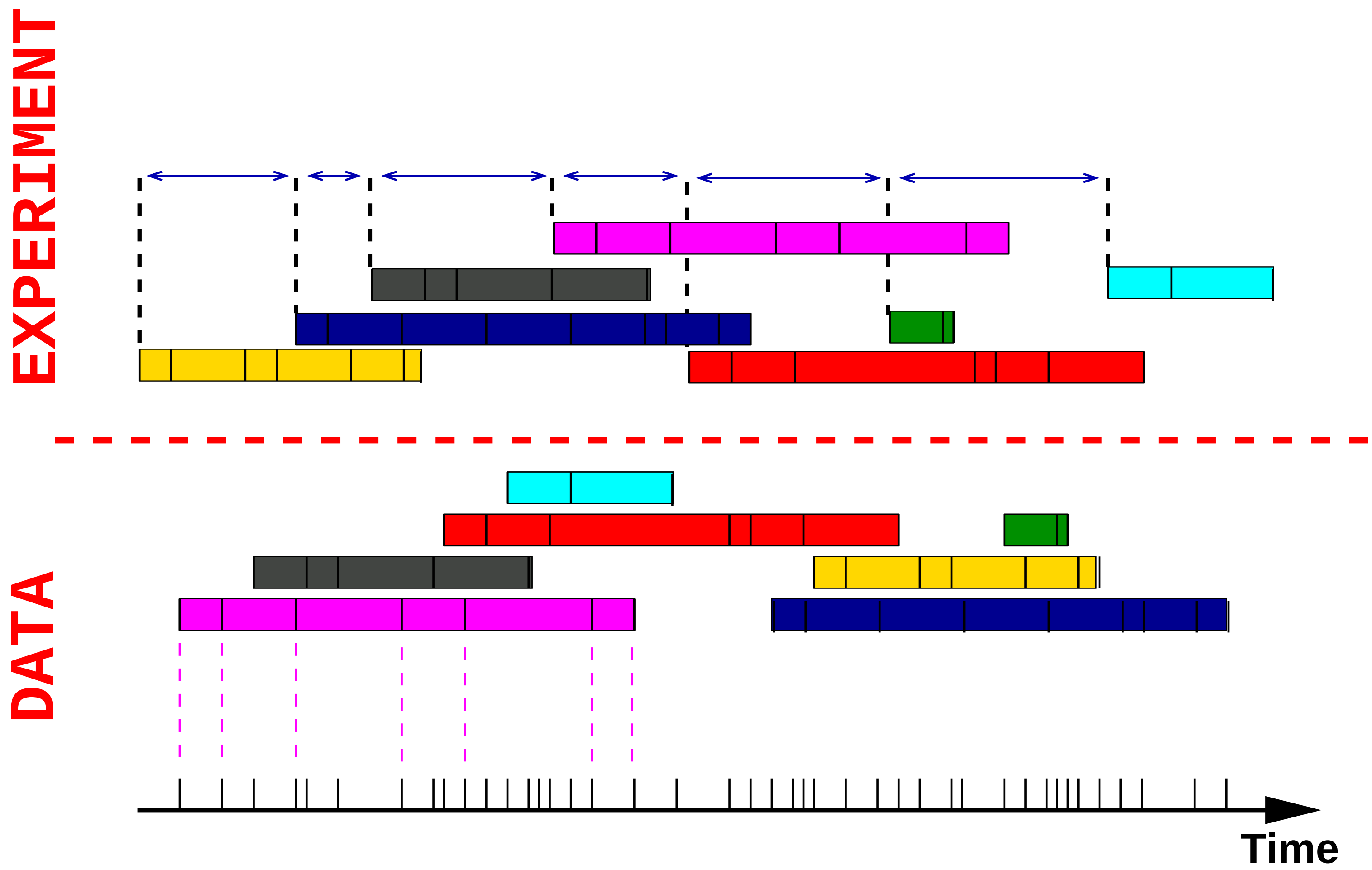
Original Order Poisson Arrivals [A-Pord]



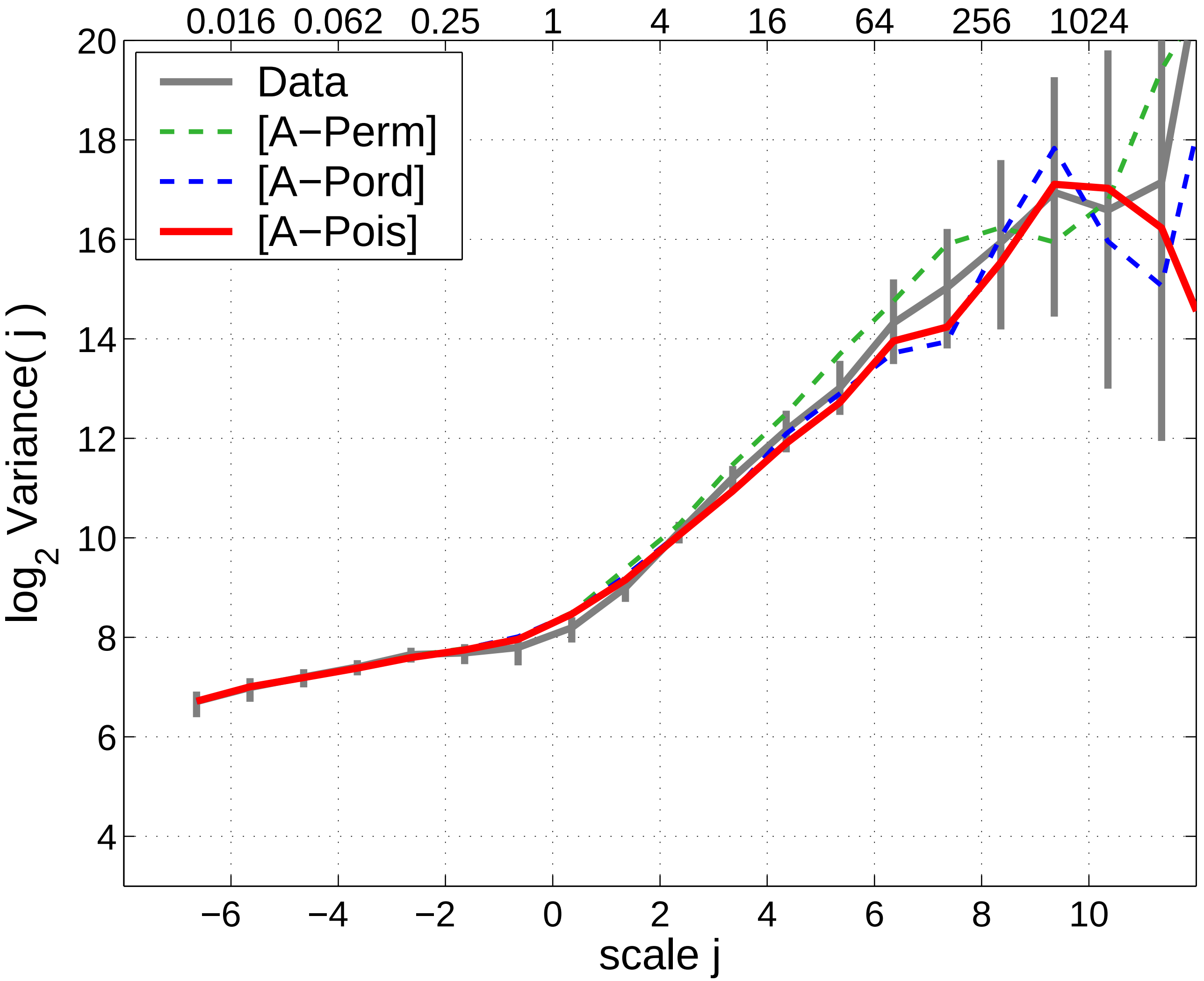
Original Order Poisson Arrivals [A-Pord]



Permuted Poisson Arrivals [A-Pois]



Permuted Poisson Arrivals [A-Pois]



From simple (Semi-)Experiments, we learn a lot

From these flow arrival *manipulations*:

- Correlations between flows can be **n**eglected
  - No need for session level hierarchical models
  - TCP dynamics between flows can be neglected
- For IP modelling, flow arrivals can be modelled as **P**oisson
  - Justifies an assumption commonly used in traffic modelling.

Note: true flow arrival process is **L**RD.

Semi-Experiment Outcomes

Flow Arrival Manipulation

- Correlation between flows can be **n**eglected,
- For IP modelling, flow arrivals can be modelled as **P**oisson,
- [A-Clus]: knee of Y affects knee of X!

Packets within Flow Manipulation

- LRD not caused by arrival process of packets within flows,
- Small scale behaviour governed by **s**tructure within flows,
- Finite Poisson process a reasonable 0-th order model,
- [P-Pois]: indistinguishable from [P-Uni],
- [P-ConstR]; [P-ScaledR]: rate acts like a scale parameter,

Flow Selection

- Observed LRD caused by heavy-tailed flow volumes,
- [S-Thin]: random thinning consistent with independent flows,
- [S-Dur]: also kills LRD (flow duration slaved to volume),
- [S-Pkt]: LRD still present even without heavy tail!

Flow Truncation Manipulation

- [T-Pkt]: also kills LRD, makes X tend to Y,

Poisson (Barlett-Lewis) Cluster Processes

Definition

- A Poisson process of seeds (flows), initiating independent groups of points (packets):

$$X(t) = \sum_i \mathcal{G}_i(t - t_{\text{F}}(i))$$

- Group: a finite renewal process with P points and inter-arrival distribution A:

$$\mathcal{G}_i(t) = \sum_{j=1}^{P(i)} \delta\left(t - \sum_{l=1}^{j-1} A_i(l)\right)$$

Parameters

- Flow arrivals: constant intensity λ_{F}
- Flow structure:
  - Packet arrivals: A, $\frac{1}{\mathbf{E}A} = \lambda_{\text{A}} < \infty$, cf: $\Phi_A(\omega)$, $\omega > 0$
  - Flow volume: P, $\mathbf{E}P = \mu_{\text{P}} < \infty$, pgf: $G_{\text{P}}(z) = \sum_{j=0}^{\infty} p_j z^j$, $|z| \leq 1$.

Fourier Spectrum

$$\Gamma_X(\nu) \;=\; \lambda_{\text{F}}\Big(\frac{\mu_{\text{P}}}{\lambda_{\text{A}}}\Gamma_{\mathcal{G}}(\nu) + \big(S_{\mathcal{G}}(\omega) + S_{\mathcal{G}}(-\omega)\big)\Big),$$

$\Gamma_{\mathcal{G}}(\nu)$: spectrum of *stationary* renewal process with inter-arrivals A, and

$$\mathcal{R}(\,S_{\mathcal{G}}(\omega)\,) = \frac{\Phi_A(\omega)}{(1-\Phi_A(\omega))^2}\Big(G_{\text{P}}(\Phi_A(\omega)) - 1\Big).$$

Verify LRD:

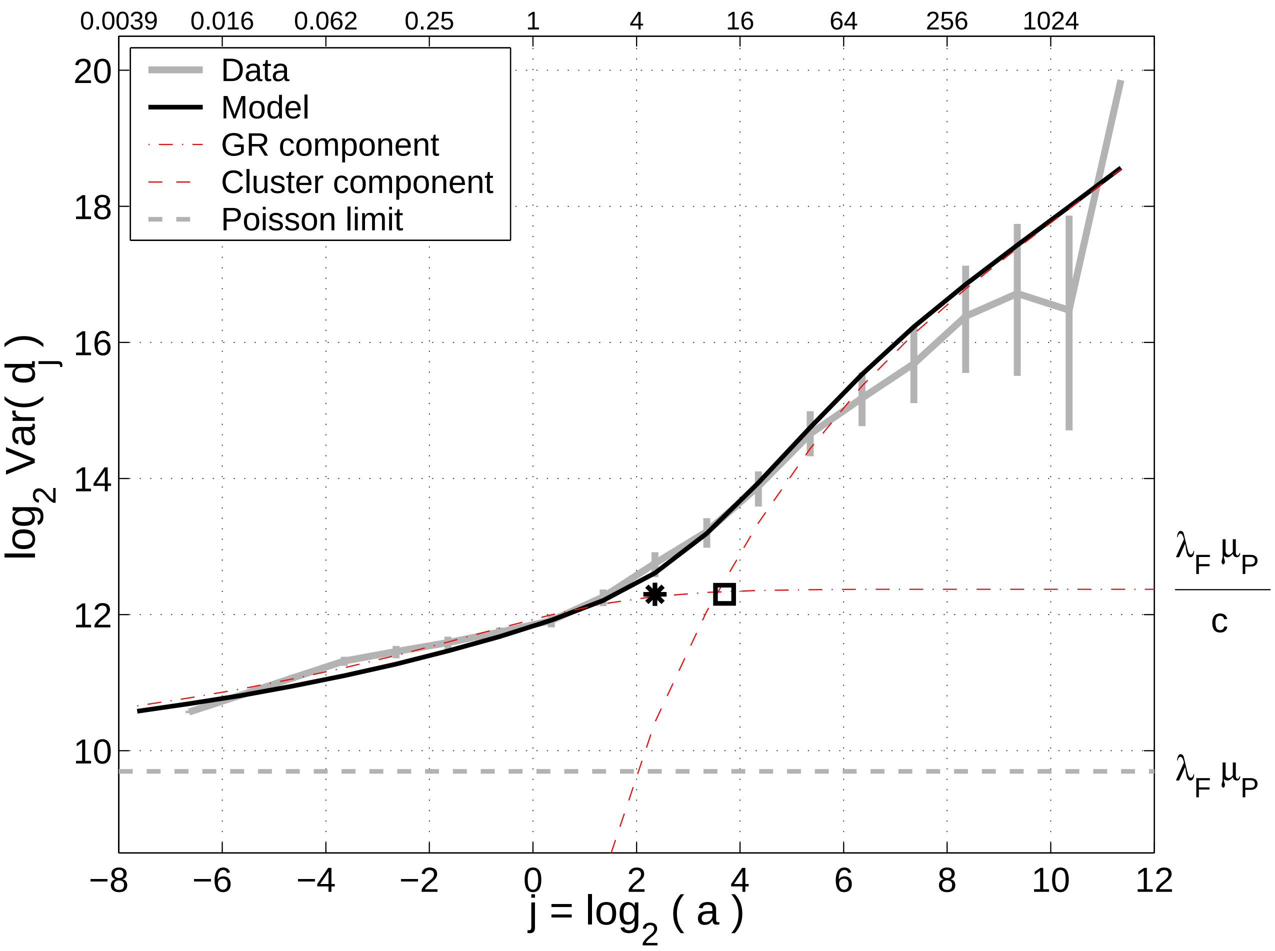
$$\begin{aligned}\mathcal{R}(\,S_{\mathcal{G}}(\omega)\,) \quad &\overset{\omega\to 0}{\sim} \quad LB(\beta)(2\pi\lambda_{\text{A}})^{2-\beta}\,\omega^{-(2-\beta)} \to \infty \\ &\overset{\omega\to\infty}{\sim} \quad -\frac{\cos(c\pi/2)}{(b\omega)^c} \to 0\end{aligned}$$

where $B(\beta) = \psi(1-\beta)\cos(\pi\beta/2)/(2\pi)^{(2-\beta)} > 0$

Properties

- λ_{F} just a variance multiplier: ‘more of same’
- has scale parameter $1/\lambda_{\text{A}}$ if A has: $\Gamma_X(\omega; \lambda_{\text{F}}, \lambda_{\text{A}}, c, F_{\text{P}}) = \Gamma_X(\omega/\lambda_{\text{A}}; \lambda_{\text{F}}, 1, c, F_{\text{P}})$
- Two terms dominating small-large scales
  - First term (small scales): scaled renewal process, no detailed P dependence
  - Second term (large scales): LRD, no A dependence beyond λ_{A}

Modelling a Typical Auckland IV Trace



$$* : \quad j^*_{\text{GR}} = -\log_2 \lambda_{\text{A}} + \log_2(\pi^2(c+1)/(3\epsilon c^2))$$

$$\square : \quad j^{**}_{\text{PGR}} = -\log_2 \lambda_{\text{A}} + \frac{1}{2-\beta}\Big(\log_2 \mu_{\text{P}} - \log_2(2LB(\beta)) - \log_2 c\Big)$$

Comparison between data and fitted Cluster Model